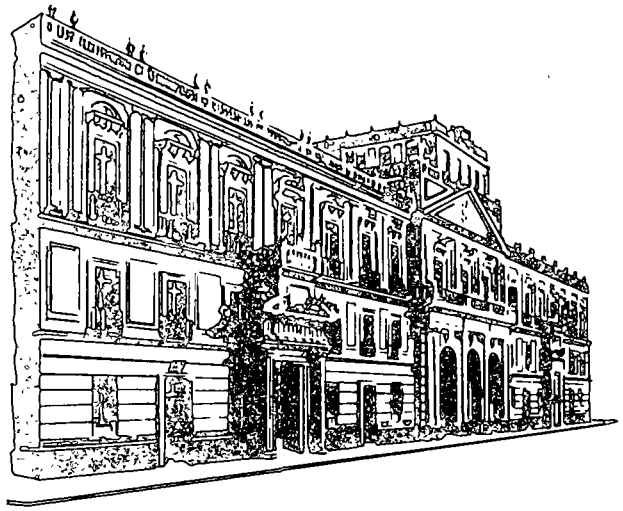


INSCRIPCIONES

CENTRO DE EDUCACION CONTINUA DE LA DIVISION DE ESTUDIOS SUPERIORES DE LA FACULTAD DE INGENIERIA, U. N. A. M.

Palacio de Mineria Calle de Tacuba No. 5
México 1, D. F.

Horario de oficinas:
Lunes a viernes de 9 a 18 h.

Cuota de inscripción $ 4,000.00

La cuota de inscripción incluye:
- una carpeta con las notas de los profesores
- bibliografía sobre el tema
- servicio de cafetería
- comidas

Requisitos
- Pagar la cuota de inscripción o traer oficio de la empresa o institución que ampare su inscripción, a más tardar una semana antes del inicio del curso
- Llenar la solicitud de inscripción

Para mayores informes hablar a los teléfonos

521-40-20 **521-73-35** **512-31-23**

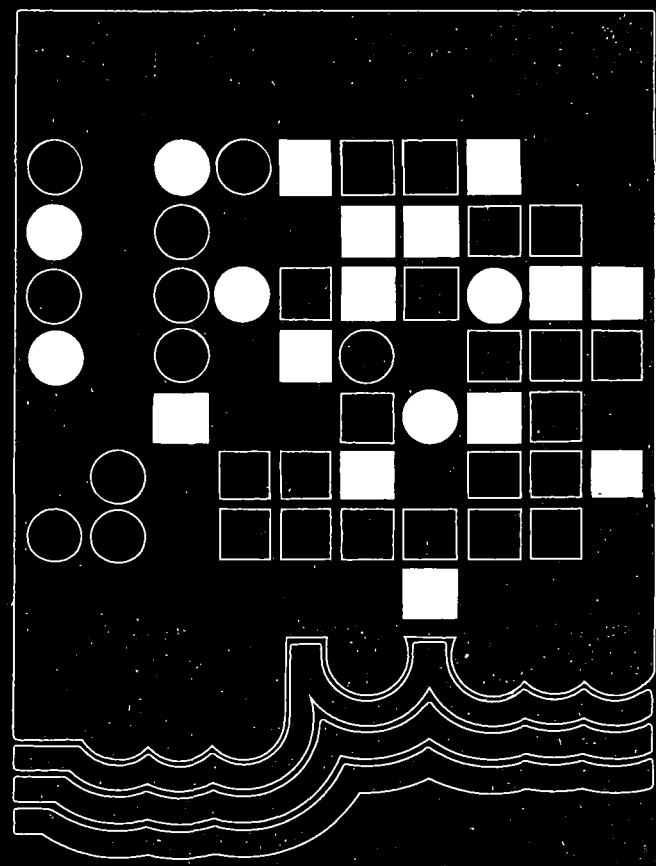
CONSTANCIA DE ASISTENCIA

Las autoridades de la Facultad de Ingeniería de la U.N.A.M., otorgarán una constancia de asistencia a los participantes que concurran regularmente y que realicen los trabajos que se les asignen durante el curso.

centro de educación continua
división de estudios superiores
facultad de ingeniería, u n a m

Palacio de Minería
Calle de Tacuba No. 5
México 1, D.F.

78



USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS

Duración: 40 h
Fechas: del 8 al 12 de Mayo
Horario: lunes a viernes de 9 a 13 h y 14 a 18 h
Coordinador: Dr. Gustavo Ayala Milian

En colaboración con el Instituto de Ingeniería, UNAM.

centro de educación continua
división de estudios superiores
facultad de ingeniería, u n a m

OBJETIVOS DEL CURSO

En la aplicación de computadoras a problemas de hidráulica en países en desarrollo, el problema no estriba en la disponibilidad de computadoras, sino en la dificultad de comunicar a los ingenieros el "saber como" involucrado en los problemas de cálculo. En este curso se proporciona el entrenamiento necesario en Hidráulica, Métodos Numéricos y Desarrollo de "Software", que permite al ingeniero la realización y-/o uso eficiente de programas de computadora aplicables a problemas de interés práctico.

A QUIEN VA DIRIGIDO

El curso ha sido diseñado para aquellas personas que trabajan en disciplinas relacionadas con la Ingeniería Hidráulica e Ingeniería Ambiental y que se enfrentan de alguna manera a problemas de dispersión de contaminantes y movimientos en grandes volúmenes de agua.

TEMARIO

1.- INTRODUCCION.

2.- DESARROLLOS RECIENTES EN EL USO DE COMPUTADORES EN PROBLEMAS DE HIDRAULICA.

3.- METODOS DE LAS DIFERENCIAS FINITAS.

4.- METODOS DE LOS ELEMENTOS FINITOS.

5.- PROPIEDADES FISICAS DE LAGOS Y AGUAS COSTERAS.

6.- CIRCULACION ESTACIONARIA EN CUERPOS DE AGUA HOMOGENEOS.

7.- APLICACION DE METODOS NUMERICOS.

8.- CIRCULACION TRANSITORIA.

9.- DISPERSION DE CONTAMINANTES

10.- CONSIDERACION DE ESTRATIFICACION.

11.- DESCRIPCION Y USO DE LOS PROGRAMAS DE COMPUTADORA.

PROFESORES

DR. GUSTAVO AYALA MILIAN

DR. GERARDO HIRIART

DR. PEDRO MARTINEZ PEREDA

CONFERENCISTAS INVITADOS

DR. JAMES A. LIGGETT
Profesor de la Universidad de Cornell

DR. JEROME J. CONNOR
Profesor del Instituto Tecnológico de Massachusets.

NOTA:

Los cursos tienen cupo limitado.

Es recomendable inscribirse con oportunidad pára garantizar su asistencia.

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

A LOS ASISTENTES A LOS CURSOS DEL CENTRO DE EDUCACION CONTINUA

Las autoridades de la Facultad de Ingeniería, por conducto del Jefe del Centro de Educación Continua, otorgan una constancia de asistencia a - quienes cumplan con los requisitos establecidos para cada curso. Las personas que deseen que aparezca su título profesional precediendo a - su nombre en la constancia, deberán entregar copia del mismo o de su - cédula a más tardar el SEGUNDO DIA de clases, en las oficinas del Centro con la señorita encargada de inscripciones.

El control de asistencia se llevará a cabo a través de la persona encargada de entregar las notas del curso. Las inasistencias serán computadas por las autoridades del Centro, con el fin de entregarle constancia solamente a los alumnos que tengan un mínimo del 80% de asistencia.

Se recomienda a los asistentes participar activamente con sus ideas y experiencias, pues los cursos que ofrece el Centro están planeados para que los profesores expongan una tesis, pero sobre todo, para que coordinen las opiniones de todos los interesados constituyendo verdaderos seminarios.

Es muy importante que todos los asistentes llenen y entregen su hoja - de inscripción al inicio del curso. Las personas comisionadas por alguna institución deberán pasar a inscribirse en las oficinas del Centro en la misma forma que los demás asistentes, entregando el oficio respectivo.

Con objeto de mejorar los servicios que el Centro de Educación Continua ofrece, al final del curso se hará una evaluación a tráves de un cuestionario diseñado para emitir juicios anónimos por parte de los asistentes.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE INGENIERIA
DIVISION DE ESTUDIOS SUPERIORES
CENTRO DE EDUCACION CONTINUA
DIRECTORIO GENERAL

REGISTRO DE ASISTENTES Y PROFESORES.

NOMBRE DEL CURSO: ______________ FOLIO [ ] 1 5 CLAVE ASOC [ ] 6 7

[ ] ※ HUMBERTO FALCON HERRERA
8 13 14 NOMBRE(S) APELLIDO PATERNO APELLIDO MATERNO 41

REG. FED. CAUS. FAHH480325 (42–51) CED. PROF. 232780 (52–58)

TEL. PARTICULAR 3929916 (59–65) TEL. OFICINA 5192770 (66–72) EXTENSION [ ] (73–76)

MARQUE CON UNA CRUZ

ASISTENTE [X] PROFESOR [ ] [ ] 77 [1] 80

AND 37 TEMOLUCO EDIF 17 DPTO 305-C
8 DOMICILIO PARTICULAR (CALLE. NUMERO Y No. INTERIOR) 41

ACUEDUCTO DE GUADALUPE Z.P. 14
42 COLONIA 71 72 73

D.F. [ ] 74 75
ESTADO

INGENIERO CIVIL [ ] 76 77 ______________ ESPECIALIDAD [ ] 78 79 [2] 80
TITULO PROFESIONAL

[ ]
8 DOMICILIO DE OFICINA (CALLE, NUMERO Y No. INTERIOR) 41

[ ] Z.P. [ ]
42 COLONIA 71 72 73

______________ [ ] 74 75 [3] 80
ESTADO

ASOCIACIONES A LAS QUE PERTENECE.

______________ [ ] ______________ [ ]

______________ [ ] ______________ [ ]

FCEC-01

SIN VELOCIDAD

CON VELOCIDAD

8/MAYO/78

1) LAGUNA DE CUITZEO

HOMOGENEO

ESTRATIFICADO

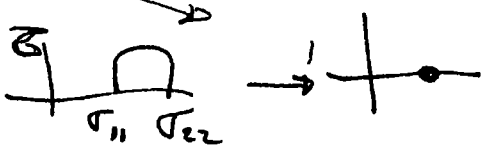
FLUIDOS IDEALES O PERFECTO: Es aquel que no puede trasmitir esfuerzos cortantes

$$\sigma_{ij} = 0 \qquad i \neq j$$

$$\left.\begin{matrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\end{matrix}\right\} = \sigma$$ puesto que no hay cortante.

$$\sigma = -p$$

$$\sigma_{ij} = -p\,\delta_{ij}$$



∇

$\vec{g} = \rho \vec{g}$

ECUACION DE ESTADO $\rho = (p, T)$

$\rho = c p$ (c te)

$$\rho\left(v_i \frac{\partial v_i}{\partial x_j} + \frac{\partial v_i}{\partial t}\right) = \frac{\partial \sigma_{ij}}{\partial x_j} + \rho \cdot F_i$$

Fluidos Viscosos $\quad \sigma_{ij} = C_{ijkl}\,\dot{\epsilon}_{re}$

$$\sigma_{ij} = \lambda \dot{\epsilon}_{kk}\delta_{ij} + 2\mu\,\dot{\epsilon}_{ij}$$

$$\sigma_{ij} = -p\delta_{ij} + \lambda\,\dot{\epsilon}_{kk}\delta_{ij} + 2\mu\,\dot{\epsilon}_{ij}$$

condicion de stokes $\quad \frac{1}{3}\sigma_{ii} = -p$

$$\sigma_{ii} = -3p + (3\lambda + 2\mu)\,\dot{\epsilon}_{kk}$$

modulo de viscocidad volumetrica $\quad K = 3\lambda + 2\mu$

$$3\lambda = -2\mu$$

$$\sigma_{ij} = -p\delta_{ij} * -\left(\tfrac{2}{3}\right.$$

$$U = \frac{\mu}{\rho}$$

Si es incompresible $\quad \sigma_{ij} = -p\delta_{ij} + 2\mu\,\dot{\epsilon}_{ij}$

$$\rho\left(\frac{\partial v_i}{\partial t} + v_j\frac{\partial v_i}{\partial x}\right) = \rho F_i + \partial\frac{\sigma_{ij}}{\partial x_j}$$

$$" \qquad = \rho F_i - \frac{\partial p}{\partial x_i} + 2\mu\frac{\partial\,\dot{\epsilon}_{ij}}{\partial x_j}$$

Shaums. Outlines Series

INTRODUCTION TO the Mechanics
of Continuous Media

Prentice Hall. 1969

Malvern L.E.

DIRECTORIO DE PROFESORES
USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS

DR. GUSTAVO AYALA MILIAN
Investigador
Instituto de Ingeniería, UNAM
Ciudad Universitaria
México 20, D.F.
Tel: 548. 97.94

DR. GERARDO HIRIART
Coordinador de la Sección de
Fluidos y Térmica
Instituto de Ingeniería, UNAM
Ciudad Universitaria
México 20, D.F.
Tel: 548. 07. 35

DR. PEDRO MARTINEZ PEREDA
Jefe del Centro de Educación Continua
D.E.S.F.I., UNAM
Palacio de Minería
Tacuba No. 5 - 1° piso
México 1, D.F.
Tel: 512.13.57

DR. JAMES A. LIGGETT
Profesor
Universidad de Cornell
U.S.A.

DR. JEROME J. CONNOR
Profesor
Instituto Tecnológico de Massachusets.
U.S.A.

'pmc.

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION
Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS

INTRODUCCION A LA MECANICA DE LOS MEDIOS CONTINUOS

MAYO, 1978.

Palacio de Minería Calle de Tacuba 5, primer piso México 1, D. F. Tel.: 521-40-20 5 Líneas

1. GOVERNING EQUATIONS - FLUID

1.1 Coordinate systems; particle derivative

We select a fixed orthogonal reference frame having directions X_1, X_2, X_3 as shown in Figure 1-1. The coordinates of a point at time t are x_1, x_2, x_3, and the coordinates at time t = 0 are a_1, a_2, a_3. In a "Lagrangean" formulation, one takes the initial coordinates (a_i) and time (t) as the independent variables

$$\text{LAGRANGE} \Rightarrow \quad x_i = x_i(a_1, a_2, a_3, t) \tag{1-1}$$

This is quite reasonable for a solid since the change in shape of the body is small. However, a fluid undergoes significant deformation and it is more convenient to take the coordinates at time t as the independent variables. This choice is called a "Eulerian" formulation.

$$\text{EULER} \Rightarrow \quad a_i = a_i(x_1, x_2, x_3, t) \tag{1-2}$$

All dependent variables such as pressure, temperature, velocity, etc., are considered to be functions of (x_1, x_2, x_3, t) in an Eulerian formulation.

Let us consider a scalar function, $f(x_i, t)$, associated with a particle (see Figure 1-2). We can express the total change in f due to change in position of the particle and time increment as:

$$\Delta f = \delta f + \tfrac{1}{2}\delta^2 f + \ldots \tag{1-3}$$

where δf is the "first" order change,

$$\delta f = \frac{\partial f}{\partial x_1}\Delta x_1 + \frac{\partial f}{\partial x_2}\Delta x_2 + \frac{\partial f}{\partial x_3}\Delta x_3 + \frac{\partial f}{\partial t}\Delta t \tag{1-4}$$

The limit, as $\Delta t \to 0$, of $\delta f/\Delta t$ is defined as the "particle" or "Stokes" derivative and is written as Df/Dt.

$$\frac{Df}{Dt} = \lim_{\Delta t \to 0} \frac{\delta f}{\Delta t} = \frac{\partial f}{\partial t} + \sum \lim_{\Delta t \to 0} \frac{\Delta x_i}{\Delta t} \cdot \frac{\partial f}{\partial x_i} \tag{1-5}$$

Now, the velocity vector for a particle is defined as

$$\vec{v} = \text{velocity vector} = \lim_{\Delta t \to 0} \frac{\vec{\Delta r}}{\Delta t} = \frac{\vec{Dr}}{Dt} \tag{1-6}$$

Operating on $\vec{r}$,

$$\vec{r} = x_1 \vec{\ell}_1 + x_2 \vec{\ell}_2 + x_3 \vec{\ell}_3 \tag{1-7}$$

and noting that $\partial x_i/\partial t = 0$ since x_i are independent variables, we obtain

$$\frac{\vec{Dr}}{Dt} = \left(\lim_{t \to 0} \frac{\Delta x_j}{\Delta t} \right) \vec{\ell}_j = v_j \vec{\ell}_j \tag{1-8}$$

v_j = component of $\vec{v}$ in the x_j direction $= \dfrac{Dx_j}{Dt}$

Finally, we can write (1-5) as :

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + \sum \frac{\partial f}{\partial x_j} v_j \tag{1-9}$$

The first term is the "local" rate of change and the remaining terms are "convective" terms.

In the Lagrange representation, one writes

$$x_i = a_i + a_i(a_1, a_2, a_3, t) \tag{1-10}$$

where u_i is the displacement from the initial position. The velocity components reduce to

$$v_i = \lim_{\Delta t \to 0} \frac{\Delta x_i}{\Delta t} \Rightarrow \frac{\partial u_i}{\partial t} \tag{1-11}$$

since a_i are constant. Also,

$$\frac{Df}{Dt} \Rightarrow \frac{\partial f}{\partial t} \tag{1-12}$$

1.2 Equilibrium equation

We consider a specific volume at time t as shown in Figure 1-3. The external actions are represented by a distributed surface loading $(\vec{p})$ and a distributed body force $(\vec{b})$. Applying Newton's law leads to the following equilibrium conditions for the total volume,

$$\iiint \vec{b}\,dV + \iint \vec{p}\,dS = \iiint \rho \frac{D\vec{v}}{Dt}\,dV \tag{1-13}$$

$$\iiint (\vec{r} \times \vec{b})dV + \iint (\vec{r} \times \vec{p})dS = \iiint \rho(\vec{r} \times \frac{D\vec{v}}{Dt})dV \tag{1-14}$$

where the integrals pertain to the position at time t; ρ is the mass density; and $\vec{Dv}/Dt$ is the acceleration vector.

We obtain "micro" equilibrium equations by expanding the surface integral involving $\vec{p}$ in terms of stress vectors and then applying Gauss's integration by parts formula. We define $\vec{\sigma}_j$ as the stress vector (force per unit area) acting on the $+j$ face, i.e. the face whose outward normal points in the $+ X_j$ direction as shown in Figure 1-3(b). The cartesian component representation is

$$\vec{\sigma}_j = \sigma_{jk} \vec{\ell}_k$$

$$\sigma_{jk} = \sigma_{jk} (x_1, x_2, x_3, t) \tag{1-15}$$

We are using the indicial summation convention here for convenience. The stress vector acting on the "n" face is given by (see figure 1-3(c))

$$\vec{\sigma}_n = \alpha_{n1} \vec{\sigma}_1 + \alpha_{n2} \vec{\sigma}_2 + \alpha_{n3} \vec{\sigma}_3 = \alpha_{nj} \cdot \vec{\sigma}_j$$

$$\alpha_{nj} = \cos (n, X_j) \tag{1-16}$$

where n is the "outward" normal. Finally, the components in the n and s directions are

$$\sigma_{nn} = \alpha_{nj} \alpha_{nk} \sigma_{jk} = \vec{\sigma}_n \vec{\ell}_n$$

$$\sigma_{ns} = \alpha_{nj} \alpha_{sk} \sigma_{jk} = \vec{\sigma}_n \cdot \vec{\ell}_s \tag{1-17}$$

where s is orthogonal to n but otherwise arbitrary.

Now, at the boundary,

$$\vec{p} = \vec{\sigma}_n = \alpha_{nj} \vec{\sigma}_j \tag{1-18}$$

Then,

$$\iint \vec{p} \, dS = \iint \alpha_{nj} \vec{\sigma}_j \, dS \tag{1-19}$$

Applying Gauss's formula,

$$\iiint g \frac{\partial f}{\partial x_i} dV = \iint \alpha_{ni} gf dS - \iiint f \frac{\partial g}{\partial x_i} dV \tag{1-20}$$

to the surface integral, we obtain finally

$$\iint \vec{p}\, dS = \iiint \frac{\partial \vec{\sigma}_j}{\partial x_j}\, dV \qquad (1\text{-}21)$$

The force equilibrium equation, (1-13), transforms to

$$\iiint \left(\vec{b} + \frac{\partial \vec{\sigma}_j}{\partial x_j} - \rho \frac{D\vec{v}}{Dt} \right) dV = \vec{0} \qquad (a)$$

and it follows that

$$\vec{b} + \frac{\partial \vec{\sigma}_j}{\partial x_j} = \rho \frac{D\vec{v}}{Dt} \qquad \text{in } V \qquad (1\text{-}22)$$

Substituting for $\vec{b}$, $\vec{p}$ in (1-14), the moment equilibrium equation, leads to

$$\vec{\ell}_j \times \vec{\sigma}_j = \vec{0} \qquad \text{in } V \qquad (1\text{-}23)$$

The scalar equations are:

$$\left.\begin{aligned} \frac{\partial}{\partial x_j}\ \sigma_{jk} + b_k &= \rho \frac{Dv_k}{Dt} \\ \sigma_{jk} &= \sigma_{kj} \end{aligned}\right\} \quad \text{in } V$$

$$p_j = \alpha_{nk}\, \sigma_{kj} \qquad \text{on } S \qquad (1\text{-}24)$$

where

$$\frac{Dv_k}{Dt} = \frac{\partial v_k}{\partial t} + v_i \frac{\partial v_k}{\partial x_i}$$

One can obtain an alternate form of the "macro" equilibrium equations by integrating the right hand side of (1-13) and (1-14). Considering (1-13), the right hand term transforms to

$$\iiint \rho \frac{D\vec{v}}{Dt}\, dV = \iint \alpha_{ni}\, v_i\, \vec{\rho v} dS \qquad (a)$$

$$+ \iiint \left\{ \frac{\partial}{\partial t}(\rho \vec{v}) - \vec{v}\left[\rho v_{i,i} + \frac{D\rho}{Dt} \right] \right\} dV$$

To interpret these terms, we consider the volume <u>fixed</u> in space (see Fig. 1-4). The first term is the momentum flow <u>out</u> of the domain and the second term is the local rate of change in momentum. The third term relates the outward mass flow and local rate of change in density. To show this, we expand the particle derivative,

$$\rho\, v_{i,i} + \frac{D\rho}{Dt} = \frac{\partial}{\partial x_i}(\rho v_i) + \frac{\partial \rho}{\partial t} \qquad (a)$$

The inward mass flow is equal to the rate of change in density,

$$\iint_S - \rho(\alpha_{ni}\, v_i) dS = \iiint \frac{\partial \rho}{\partial t}\, dV \qquad (b)$$

Integrating by parts leads to a "continuity" condition,

$$\iiint \left\{ \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho v_i) \right\} dV = 0$$

$$\Downarrow$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho v_i) = 0 \qquad \text{in } V \qquad (1\text{-}25)$$

With (1-25), we can write the force equilibrium equation as:

$$\iiint \vec{b} dV + \iint \vec{p} dS = \iint \vec{v}(\rho v_n) dS + \iiint \frac{\partial}{\partial t}(\rho \vec{v}) dV \qquad (1\text{-}26)$$

Eq. (1-26) is called the "momentum" equation. One should note that (1-26) + (1-25) is equivalent to the original form of the equilibrium equation, (1-13).

We could have established (1-25) by allowing the volume to move and requiring no mass flow across the boundary. In this approach, the particle derivative of the total mass is zero.

$$\frac{D}{Dt}\iiint \rho dV = \iiint \left(\frac{D\rho}{Dt} dV + \rho \frac{D}{Dt}(dV) \right) = 0 \qquad (1\text{-}27)$$

We will show later than

$$\frac{D}{Dt}(dV) = (\partial v_i/\partial x_i)dV \qquad (1\text{-}28)$$

Before moving on to kinematic relations, we comment briefly on an inviscid fluid. If the shear stress components in the stress tensor are neglected,

$$\sigma_{ij} = 0 \qquad i \neq j \qquad (1\text{-}29)$$

the fluid is called "frictionless" or inviscid. To determine whether there are any relations between the nonzero elements $(\sigma_{11}, \sigma_{22}, \sigma_{33})$, we consider the stress transformation law, (1-17),

$$\begin{aligned} \sigma_{ns} &= \alpha_{nj}\, \alpha_{sk}\, \sigma_{jk} \\ &= \alpha_{nk}\, \alpha_{sj}\, \sigma_{jj} \qquad \text{for } \sigma_{ij} = \delta_{ij}\, \sigma_{ij} \end{aligned} \qquad (a)$$

But, $\sigma_{ns} = 0$ for $s \neq n$. This requires

$$\sigma_{11} = \sigma_{22} = \sigma_{33} = \sigma \qquad (1\text{-}30)$$

and the remaining transformation law reduces to

$$\sigma_{nn} = \alpha_{nj}\, \alpha_{nk}\, \sigma_{jk} = \sigma \qquad (1\text{-}31)$$

The state of stress is defined by a single variable, σ. For a fluid,

σ is negative (compression) and therefore we shall take

$$\sigma = - p \qquad \text{(p denotes pressure)} \qquad (1\text{-}32)$$

The equilibrium equations and boundary conditions for a frictionless fluid are:

$$\begin{aligned} - \frac{\partial}{\partial x_k} p + b_k &= \rho \frac{Dv_k}{Dt} \qquad \text{in } V \\ p_n &= - p \qquad \text{on } S \end{aligned} \qquad (1\text{-}33)$$

Note that one cannot apply a tangential boundary force to an inviscid fluid, i.e. p_s must be zero.

1.3 Principle of Virtual Power

We derive the Principle of Virtual Power by operating on the stress equilibrium equations and stress-boundary force relations,

$$\begin{aligned} \frac{\partial}{\partial x_j} \sigma_{jk} + b_k &= \rho \frac{Dv_k}{Dt} \qquad \text{in } V \qquad \text{(a)} \\ p_k &= \alpha_{nj} \sigma_{jk} \qquad \text{on } S \end{aligned}$$

The "true" solution for the stress components satisfies (a). Let us multiply (a) by a function, say $\bar{v}_k$, and integrate over the domain.

$$\iiint \left\{ \frac{\partial}{\partial x_j} \sigma_{jk} + b_k - \rho \frac{Dv_k}{Dt} \right\} \bar{v}_k \, dV + \iint \{ p_k - \alpha_{nj} \sigma_{jk} \} \bar{v}_k dS \equiv 0 \qquad (b)$$

Equation (b) must be satisfied for arbitrary $\bar{v}_k$ if the stress field is an equilibrium field. Next, we integrate the stress term with Gauss's

formula, and obtain the "principle of virtual power",

$$\iiint b_k \bar{v}_k dV + \iint p_k \bar{v}_k dS = \iiint \left(\sigma_{jk} \frac{\partial \bar{v}_k}{\partial x_j} + \rho \bar{v}_k \frac{Dv_k}{Dt} \right) dV \tag{1-34}$$

The left hand term can be interpreted as virtual power if one considers $\bar{v}_k$ as a "virtual" velocity. Similarly, the right hand terms can be interpreted as the virtual time rate of change (particle derivative) of the internal deformation work and the kinetic energy. Note that (1-34) applies for a particular time, t, and arbitrary $\bar{v}_k$. We emphasize that it is just an alternate statement of equilibrium.

The principle of virtual power is the basis for finite element models in fluid mechanics. Its role is similar to that of the principle of virtual displacements in solid mechanics. If we take $\bar{v}_k = v_k$, the actual velocity, the principle of virtual power coincides with the first law of thermodynamics,

Rate of External Work = Rate of deformation work
+ rate of change of the kinetic energy

(1-35)

Also, we observe that the particle derivatives of the deformation measures are the quantities required. The total (actual) deformations do not appear in the variational statement. Finally, one can utilize the principle of virtual power to establish "consistent" boundary conditions. For example, if one assumes the fluid is inviscid,

$$\iiint \sigma_{jk} \frac{\partial}{\partial x_j} \bar{v}_k \, dV \Rightarrow \iiint - p \frac{\partial}{\partial x_k} \bar{v}_k \, dV \tag{a}$$

$$= \iint - p \bar{v}_n \, dS + \iiint \bar{v}_k \frac{\partial}{\partial x_k} p dV$$

Then, expanding the surface force term,

$$\iint p_k \bar{v}_k \, dS = \iint (p_n \bar{v}_n + p_s \bar{v}_s) dS \tag{b}$$

we obtain the "consistent" stress - surface force boundary condition

$$\begin{aligned} p_n &= -p \\ p_s &= 0 \end{aligned} \tag{c}$$

and the equilibrium equation,

$$b_k = \frac{\partial}{\partial x_k} p + \rho \frac{Dv_k}{Dt} \tag{d}$$

1.4 Kinematic Relations

Our objective in this section is to establish expressions for the time rate of change of the deformation measures. To simplify the discussion, we consider initially the 2-dimensional case. The 3-dimencional expressions can be obtained by generalising the 2-dimensional expressions.

Figure 1-5 shows the initial (time t) and deformed (time t + Δt) positions of 2 differential line elements. We visualise the movement of a line to consist of translation, rotation, and extension. Let $\Delta\varepsilon_1$ denote the relative incremental extension of line PQ.

$$\Delta\varepsilon_1 = \frac{|\overrightarrow{P'Q'}| - |\overrightarrow{PQ}|}{|\overrightarrow{PQ}|} = \frac{1}{dx_1} |\overrightarrow{P'Q'}| - 1 \tag{1-36}$$

Substituting for $|P'Q'|$, we obtain after some algebra,

$$\Delta\varepsilon_1 (1 + \tfrac{1}{2}\Delta\varepsilon_1) = v_{1,1} \Delta t + \tfrac{1}{2}\{ (v_{1,1} \Delta t)^2 + (v_{2,1} \Delta t)^2 \} \tag{e}$$

where

$$v_{1,1} = \frac{\partial}{\partial x_1} v_1 \qquad v_{2,1} = \frac{\partial}{\partial x_1} v_2$$

We define the strain rate as the particle derivative,

$$\dot{\varepsilon}_1 = \lim_{\Delta t \to o} \frac{\Delta \varepsilon_1}{\Delta t} = \frac{D}{Dt} \varepsilon_1 \tag{1-37}$$

As $\Delta t \to 0$, the nonlinear terms vanish and we are left with

$$\dot{\varepsilon}_1 = v_{1,1} \tag{1-38}$$

By analogy,

$$\Delta \varepsilon_2 = \frac{|\overrightarrow{P'R'}| - |\overrightarrow{PR}|}{|\overrightarrow{PR}|}$$

$$\dot{\varepsilon}_2 = v_{2,2} \tag{1-39}$$

Generalizing, we can write

$$\dot{\varepsilon}_i = v_{i,i} \qquad \text{(no sum)} \tag{1-40}$$

The relative incremental volume change, $\Delta \varepsilon_v$, is determined from

$$\Delta \varepsilon_v = \frac{\Delta(\text{volume})}{\text{initial volume}} = \left| \overrightarrow{P'Q'} \times \overrightarrow{P'R'} \right| \frac{1}{dx_1 dx_2} - 1$$

$$\Downarrow \tag{1-41}$$

$$\Delta \varepsilon_v = (v_{1,1} + v_{2,2})\Delta t + (v_{1,1}\, v_{2,2} - v_{1,2}\, v_{2,1})(\Delta t)^2$$

We define $\dot{\varepsilon}_v$ as the volumetric strain rate,

$$\dot{\varepsilon}_v = \lim_{\Delta t \to o} \frac{\Delta \varepsilon_v}{\Delta t} = \frac{D}{Dt} \varepsilon_v \tag{1-42}$$

$$= v_{1,1} + v_{2,2} + v_{3,3}$$

Using the notation of vector calculus,

$$\nabla = \vec{\ell}_k \frac{\partial}{\partial x_j} \tag{1-43}$$

$$\varepsilon_v = \nabla \cdot \vec{v} = \text{Div.}\ \vec{v}$$

Lastly, we consider the rotation terms θ_{12} and θ_{21}.

$$\sin\Delta\theta_{12} = \frac{v_{2,1}\,\Delta t}{1+\Delta\varepsilon_1} \tag{1-44}$$

$$\sin\Delta\theta_{21} = \frac{v_{1,2}\,\Delta t}{1+\Delta\varepsilon_2}$$

The limits are

$$\dot{\theta}_{12} = v_{2,1} \qquad \dot{\theta}_{21} = v_{1,2} \tag{1-45}$$

The shear strain rate is the sum of $\dot{\theta}_{12}$ and $\dot{\theta}_{21}$

$$\dot{\gamma}_{12} = \dot{\theta}_{12} + \dot{\theta}_{21} \tag{1-46}$$

Generalizing,

$$\dot{\gamma}_{ij} = \dot{\theta}_{ij} + \dot{\theta}_{ji} = v_{j,i} + v_{i,j} \tag{1-47}$$

It is convenient to introduce the strain rate tensor, $\dot{e}_{ij}$,

$$\dot{e}_{ij} = \tfrac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) \tag{1-48}$$

The strain measures in terms of $\dot{e}_{ij}$ are

$$\dot{\varepsilon}_i = \dot{e}_{ii} \qquad \dot{\varepsilon}_v = \sum \dot{e}_{ii}$$

$$\dot{\gamma}_{ij} = 2\,\dot{e}_{ij} \tag{1-49}$$

Let us re-examine the rotation terms, θ_{12} and θ_{21}. When there is only rigid body motion, $\dot{\gamma} = 0$ and

$$\dot{\theta}_{12} = \omega_{12} \qquad \dot{\theta}_{21} = -\omega_{12} \tag{a}$$

where ω_{12} is the angular velocity about the X_3 axis. This suggests that we take, as a measure of the angular velocity, the difference between $\dot{\theta}_{12}$ and $\dot{\theta}_{21}$:

ω_{12} = average value of angular velocity about the X_3 axis

$$= \tfrac{1}{2}(\dot{\theta}_{12} - \dot{\theta}_{21}) = \tfrac{1}{2}(v_{2,1} - v_{1,2}) \tag{b}$$

Generalising (b), we write

$$\omega_{ij} = \tfrac{1}{2}(\dot{\theta}_{ij} - \dot{\theta}_{ji}) = \tfrac{1}{2}(v_{j,i} - v_{i,j}) \tag{1-50}$$

Cyclic permutation of the subscripts gives the average velocities about the 3 axes.

$$\begin{aligned} \omega_{12} &\rightarrow \omega_3 \\ \omega_{23} &\rightarrow \omega_1 \\ \omega_{31} &\rightarrow \omega_2 \end{aligned} \tag{1-51}$$

The tensor, ω_{ij}, is skew symmetric and is called the vorticity tensor. It is not difficult to show that ω_{ij} is <u>invariant</u> for certain transformations of axes. For example,

$$\omega'_{12} = \omega_{12} \tag{1-52}$$

where ω'_{12} corresponds to the directions (X'_1, X'_2, X_3). This is actually why we can interpret it as the average rotation about X_3.

We end this section by listing two definitions.

1) A fluid is <u>incompressible</u> when the volume strain rate is zero.

$$\dot{\varepsilon}_v = v_{1,1} + v_{2,2} + v_{3,3} = 0 \tag{1-53}$$

ρ = mass density = constant

In this case, we have to determine the pressure from an equilibrium consideration.

2) The motion is <u>irrotational</u> when $\omega_{12} = \omega_{23} = \omega_{31} = 0$. This requires

$$\begin{aligned} v_{2,1} - v_{1,2} &= 0 \\ v_{3,2} - v_{2,3} &= 0 \\ v_{1,3} - v_{3,1} &= 0 \end{aligned} \tag{1-54}$$

1.5 <u>Stress-strain relations - Newtonian Fluid.</u>

We consider first a linearly elastic solid to provide us with some background. The stress-strain relations are

$$\sigma_{ij} = \lambda \varepsilon_v \delta_{ij} + 2G e_{ij} \tag{1-55}$$

where λ, G are material properties, ε_v is the volumetric strain, and e_{ij} is the strain tensor.

To distinguish between the volumetric and shear deformation modes,

we express the strain (and stress) tensors as a combination of 2 tensors which are called the "spherical" and "deviatoric" components.

$$e_{ij} = e^s_{ij} + e^d_{ij}$$

$$\sigma_{ij} = \sigma^s_{ij} + \sigma^d_{ij}$$

where

$$e^s_{ij} = \frac{1}{3}\varepsilon_v \delta_{ij} = \varepsilon_m \delta_{ij}$$

$$\varepsilon_m = \frac{1}{3}(e_{11} + e_{22} + e_{33})$$

$$e^d_{ii} = e_{ii} - \varepsilon_m$$

$$e^d_{ij} = e_{ij} \qquad i \neq j \tag{1-56}$$

and

$$\sigma^s_{ij} = \sigma_m \delta_{ij} = \frac{1}{3}(\sigma_{11} + \sigma_{22} + \sigma_{33})\delta_{ij}$$

$$\sigma^d_{ii} = \sigma_{ii} - \sigma_m$$

$$\sigma^d_{ij} = \sigma_{ij} \qquad i \neq j$$

The mean stress can be interpreted as an "equivalent" hydrostatic pressure; σ^d_{ii} reflects the deviation from the hydrostatic state; and σ^d_{ij} are the shearing stresses.

Similarly, ε_m is the "average" extension corresponding to a volume change and e^d_{ij} leads to a change in shape, i.e., shearing deformation.

The trace of the deviatoric components vanishes,

$$\sigma^d_{11} + \sigma^d_{22} + \sigma^d_{33} \equiv 0$$

$$e^d_{11} + e^d_{22} + e^d_{33} \equiv 0 \tag{1-57}$$

which shows that the deviatoric components are associated with nonuniform states.

Using (1-55), we find

$$\sigma_m = K\,\varepsilon_v \tag{1-58}$$

$$K = \text{bulk modulus} = \frac{3\lambda + 2G}{3}$$

and

$$\sigma^d_{ij} = 2G\,e^d_{ij} = 2G\left(e_{ij} - \frac{1}{3}\varepsilon_v\,\delta_{ij}\right) \tag{1-59}$$

We introduce the assumption of incompressibility by setting $K = \infty$. In this case, σ_m has to be determined from the equilibrium equations and the displacements are constrained by $\varepsilon_v = 0$.

The stress-strain relations for a Newtonian fluid are similar in form to those of the solid except that the strains are replaced with strain rates. We write

$$\sigma_{ij} = -p\delta_{ij} + \tau_{ij} \tag{1-60}$$

where p is the hydrostatic pressure and τ_{ij} are the stresses due to motion, referred to as the "viscous" stresses. The relation for τ_{ij} is taken as

$$\tau_{ij} = \lambda^* \dot{\varepsilon}_v\,\delta_{ij} + 2\mu\,\dot{e}_{ij} \tag{1-61}$$

where λ^*, μ are viscosity coefficients. In this approach, the mean stress is determined from

$$\sigma_m = \frac{1}{3}(\sigma_{11} + \sigma_{22} + \sigma_{33})$$

$$= - p + \frac{1}{3}(\tau_{11} + \tau_{22} + \tau_{33}) \qquad (1\text{-}62)$$

$$= - p + K^* \dot{\varepsilon}_v$$

$$K^* = \frac{3\lambda^* + 2\mu}{3} = \text{coefficient of bulk viscosity}$$

$$p = p \text{ (density, temperature)} = p(\rho, T)$$

$$\frac{D\rho}{Dt} = - \rho \, \dot{\varepsilon}_v$$

The deviatoric stresses are given by

$$\sigma^d_{ij} = \sigma_{ij} - \sigma_m \delta_{ij} = 2\mu(\dot{e}_{ij} - \frac{1}{3} \dot{\varepsilon}_v \delta_{ij}) \qquad (1\text{-}63)$$

Setting $K^* = 0$ is known as the "Stokes" condition and leads to

$$\lambda^* = - \frac{2}{3} \mu$$

$$\sigma_m = \frac{1}{3}(\sigma_{11} + \sigma_{22} + \sigma_{33}) \equiv -p \qquad (1\text{-}64)$$

$$\sigma_{ij} = - p\delta_{ij} + 2\mu(\dot{e}_{ij} - \frac{1}{3} \dot{\varepsilon}_v \delta_{ij})$$

$$p = p(\rho, T)$$

With this assumption, the mean stress coincides with the thermodynamic pressure, p. We point out that (1-60) and (1-64) are valid only for laminar flow.

1.6 Summary of Governing Equations

At this point, we summarize the governing equations:

Equilibrium

$$\frac{\partial}{\partial x_j}\sigma_{jk} + b_k = -\rho\frac{Dv_k}{Dt} \qquad \text{in } V$$

$$p_j = \alpha_{nk}\,\sigma_{kj} \qquad \text{on } S \tag{1-65}$$

$$\sigma_{jk} = \sigma_{kj}$$

Kinematic

$$\dot{e}_{ij} = \tfrac{1}{2}(v_{i,j} + v_{j,i}) = \text{strain rate tensor}$$

$$\dot{\varepsilon}_i = \dot{e}_{ii} \qquad \dot{\gamma}_{ij} = 2\dot{e}_{ij}$$

$$\dot{\varepsilon}_v = \dot{e}_{11} + \dot{e}_{22} + \dot{e}_{33} \tag{1-66}$$

$$\omega_{ij} = \tfrac{1}{2}(v_{j,i} - v_{i,j}) = \text{vorticity tensor}$$

Stress-strain - Newtonian fluid

$$\sigma_{ij} = -p\delta_{ij} + \tau_{ij}$$

$$p = p(\rho,T) \tag{1-67}$$

$$\frac{D\rho}{Dt} = -\rho\,\dot{\varepsilon}_v$$

$$\tau_{ij} = \lambda^{*}\,\dot{\varepsilon}_v\delta_{ij} + 2\mu\dot{e}_{ij}$$

$$\lambda^{*} = -\frac{2}{3}\mu \qquad \text{for "Stokes" condition}$$

It is also of interest to express the principle of virtual power (1-34) in terms of the strain rate measures.

We let

$$\bar{v}_k = \delta v_k \tag{1-68}$$

and note that since the stress components are symmetrical in the subscripts,

$$\begin{aligned} \sigma_{jk} \frac{\partial}{\partial x_j} \bar{v}_k &= \tfrac{1}{2} \sigma_{jk} (\delta v_{k,j} + \delta v_{j,k}) \\ &= \sigma_{jk} \, \delta e_{jk} \end{aligned} \tag{a}$$

Then, we can write the principle of virtual power as

$$\begin{aligned} \iint p_k \, \delta v_k dS + \iiint b_k \, \delta v_k dV &= \iiint \{\sigma_{jk} \dot{\delta e}_{jk} + \rho \frac{Dv_k}{Dt} \delta v_k\} \, dV \\ &= \iiint \{- p \dot{\delta \varepsilon}_v + \tau_{jk} \dot{\delta e}_{jk} + \rho \frac{Dv_k}{Dt} \delta v_k\} \, dV \end{aligned} \tag{1-69}$$

for arbitrary δv_k

1.7. Navier Stokes equations - incompressible Newtonian fluid.

The equilibrium equations and expressions for the surface forces, (1-65), in terms of p and τ_{ij} are

$$\left.\begin{aligned} - \frac{\partial}{\partial x_k} p + \frac{\partial}{\partial x_j} \tau_{jk} + b_k &= \rho \frac{Dv_k}{Dt} \quad \text{in } V \\ p_n = - p + \alpha_{mj} \alpha_{nk} \tau_{kj} \\ p_s = \alpha_{sj} \alpha_{nk} \tau_{kj} \end{aligned}\right\} \text{ on } S \tag{1-70}$$

We take τ_{ij} according to (1-67) and constrain the velocities to

satisfy $\dot{\epsilon}_v = 0$,

$$\dot{\epsilon}_v = 0 \Rightarrow v_{1,1} + v_{2,2} + v_{3,3} = 0 \tag{1-71}$$

$$\tau_{jk} \Rightarrow 2\mu e_{jk} = \mu(v_{j,k} + v_{k,j}) \tag{1-72}$$

Also, we consider the body forces to be due to gravity, and write

$$b_k = -\rho g_k \tag{1-73}$$

Substituting for τ_{jk} and b_k in (1-70) results in the "Navier-Stokes" equations,

$$\left.\begin{aligned} \frac{\partial}{\partial x_k}(p/\rho) + g_k &= \nu v_{k,jj} - \frac{Dv_k}{Dt} \\ v_{i,i} &= 0 \end{aligned}\right\} \text{ in } V \tag{1-74}$$

and

$$\left.\begin{aligned} p_n/\rho &= -(p/\rho) + 2\nu\frac{\partial v_n}{\partial n} \\ p_s/\rho &= \nu\left(\frac{\partial v_s}{\partial n} + \frac{\partial v_n}{\partial s}\right) \end{aligned}\right\} \text{ on } S \tag{1-75}$$

where $\nu = \mu/\rho$ is the kinematic viscosity. The expression for p_s applies when the boundary is straight. For a curved boundary, we must use

$$p_s/\rho = \nu\left\{\frac{\partial}{\partial n} v_s + \frac{\partial}{\partial s} v_n - v_k \frac{\partial}{\partial k}\ \alpha_{nk}\right\} \tag{1-76}$$

The unknown variables are the velocity components, v_j, and the pressure, p. They have to satisfy the equilibrium equations and

incompressibility condition in V and the boundary condition

$$\left.\begin{matrix} p_n \text{ or } v_n \\ p_s \text{ or } v_s \end{matrix}\right\} \quad \text{prescribed on } S \tag{1-77}$$

Let us now see how we have to modify the principle of virtual power, (1-69). If we set $\dot{\varepsilon}_v = 0$, the pressure term drops out. However, the virtual velocities are no longer arbitrary but are now constrained by

$$\delta\dot{\varepsilon}_v = \delta v_{1,1} + \delta v_{2,2} + \delta v_{3,3} = 0 \tag{1-78}$$

We can include the constraint condition by introducing a Lagrange multiplier λ and requiring

$$\delta\left[\lambda\,\dot{\varepsilon}_v\right] = \lambda\delta\dot{\varepsilon}_v + \delta\lambda\dot{\varepsilon}_v = 0 \quad \text{in } V \tag{1-79}$$

for arbitrary λ. Since ε_v is volumetric strain, the physical significance of λ is hydrostatic tensile stress. Then, we reverse the sign and take $\lambda = -p$. The final form of the principle of virtual power is

$$\frac{1}{\rho}\iint(p_n\delta v_n + p_s\delta v_s)dS - \iiint g_k\delta v_k\,dV$$
$$= \iiint\left\{-\frac{p}{\rho}\delta\dot{\varepsilon}_v - \frac{\delta p}{\rho}\dot{\varepsilon}_v + \frac{1}{\rho}\tau_{jk}\delta\dot{e}_{jk} + \frac{Dv_k}{Dt}\delta v_k\right\} = 0 \tag{1-80}$$

for arbitrary δp, δv_k

Substituting

$$\frac{1}{\rho}\tau_{jk} = \nu(v_{j,k} + v_{k,j}) \qquad (a)$$

$$\dot{\delta\varepsilon}_v = \delta v_{k,k} \qquad \dot{\delta e}_{jk} = \tfrac{1}{2}(\delta v_{j,k} + \delta v_{k,j})$$

and integrating by parts leads to (1-74) and (1-75).

Finally, we introduce the assumption of frictionless flow by setting $\tau_{jk} = 0$. The reduced equations are

$$\left.\begin{aligned} -\frac{\partial}{\partial x_k}(p/\rho) + \mathcal{E}_k + \frac{Dv_k}{Dt} &= 0 \\ v_{i,i} &= 0 \end{aligned}\right\} \text{ in } V$$

$$\left.\begin{aligned} p_n &= p \quad \text{or } v = v_n \\ p_s &= 0 \quad v_s \text{ arbitrary} \end{aligned}\right\} \text{ on } S \qquad (1\text{-}81)$$

and

$$\frac{1}{\rho}\iint p_n \delta v_n \, dS = \iiint \left\{ \left[\mathcal{E}_k + \frac{Dv_k}{Dt} \right] \delta v_k - \frac{1}{\rho}\left[p\dot{\delta\varepsilon}_v + \dot{\varepsilon}_v \delta p \right] \right\} dV$$

for arbitrary δv_k, δp (1-82)

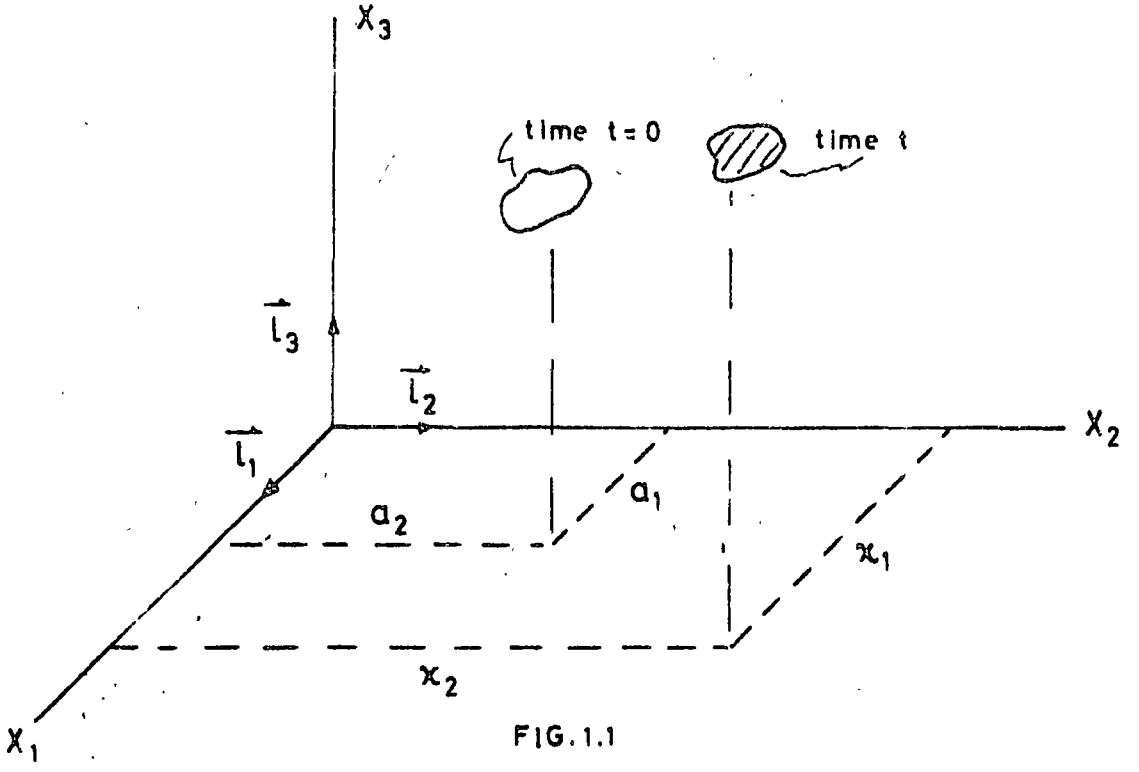
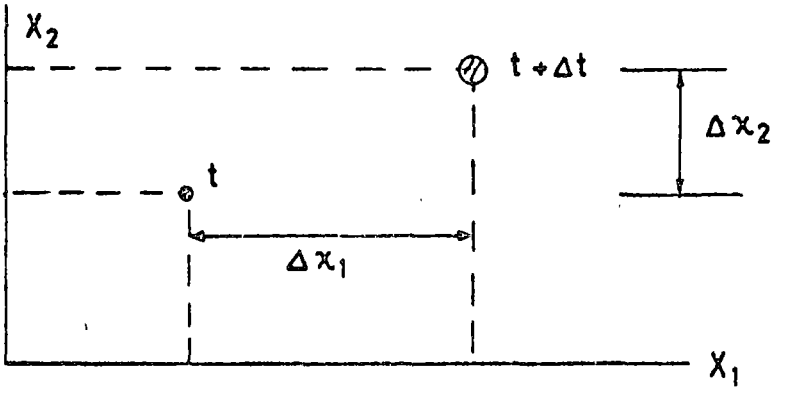

X_3
time t = 0
time t
$\vec{l}_3$
$\vec{l}_2$
$\vec{l}_1$
X_2
a_2
a_1
x_1
x_2
X_1

FIG. 1.1


X_2
$t + \Delta t$
Δx_2
t
Δx_1
X_1

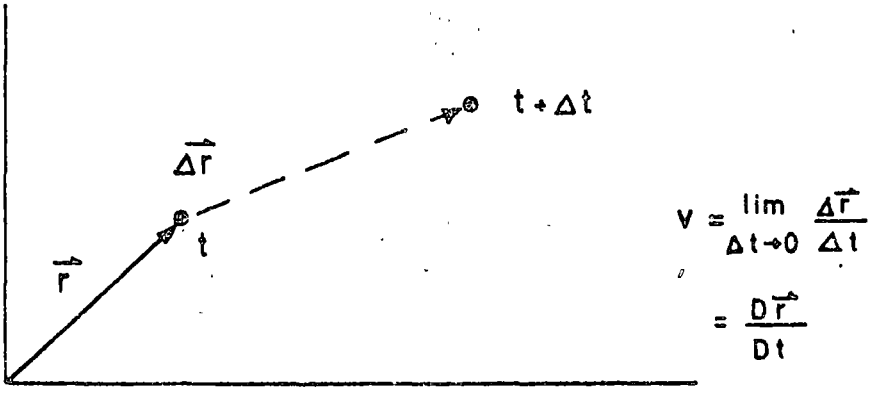

$t + \Delta t$
$\Delta \vec{r}$
t
$\vec{r}$
$v = \lim_{\Delta t \to 0} \frac{\Delta \vec{r}}{\Delta t}$
$= \frac{D\vec{r}}{Dt}$

FIG. 1.2

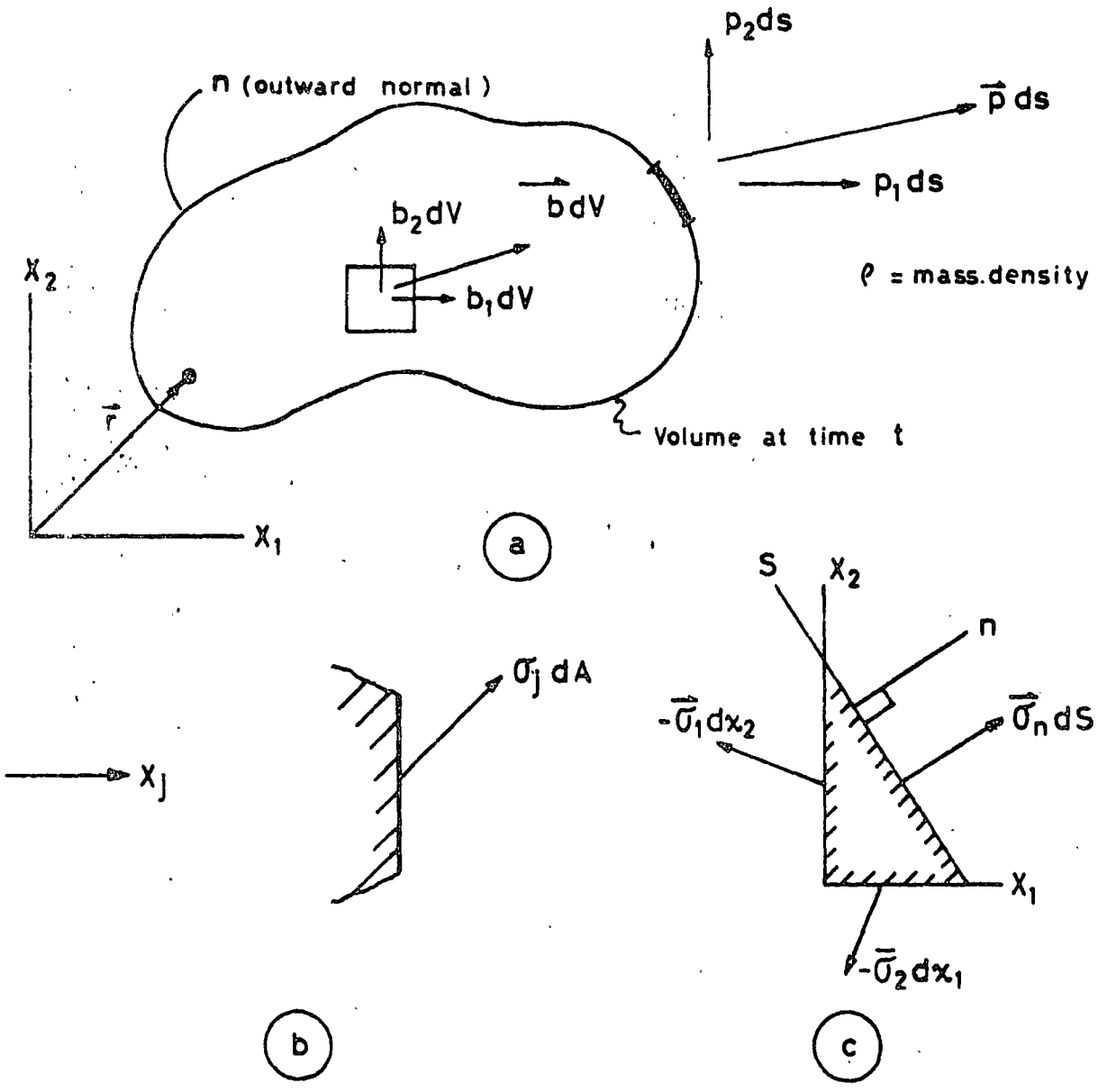

n (outward normal)
$p_2 ds$
$\vec{p}\, ds$
$p_1 ds$
$b_2 dV$
$\vec{b} dV$
$b_1 dV$
ρ = mass.density
X_2
$\vec{r}$
Volume at time t
X_1
a
$\sigma_j dA$
X_j
S
X_2
n
$-\vec{\sigma}_1 dx_2$
$\vec{\sigma}_n dS$
X_1
$-\vec{\sigma}_2 dx_1$
b
c

FIG. 1.3

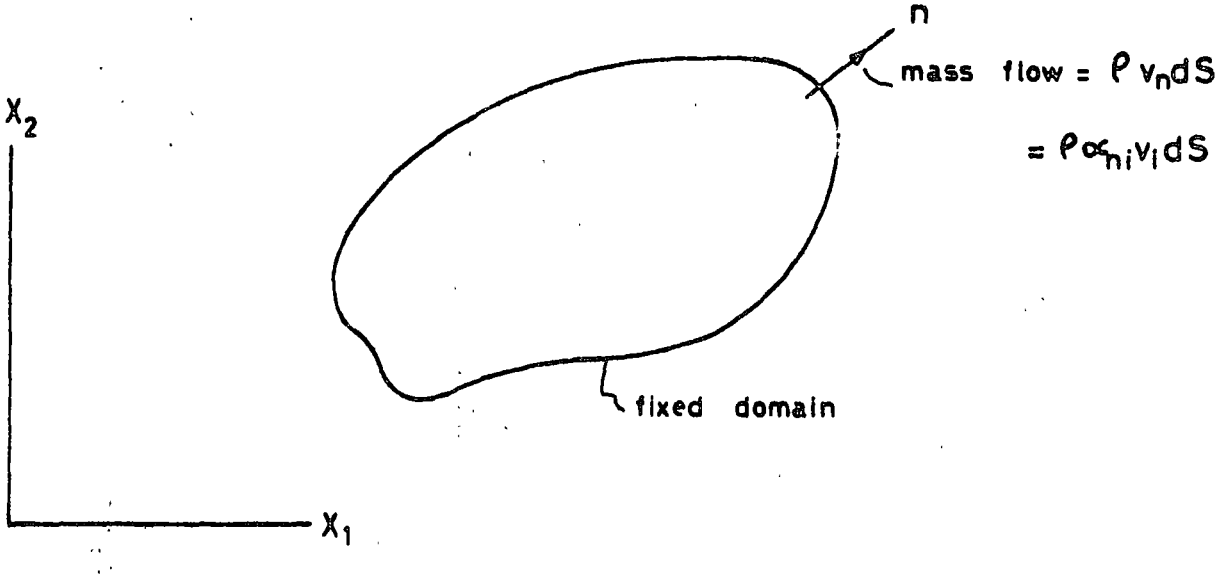

n
mass flow = $\rho v_n dS$
$= \rho \alpha_{ni} v_i dS$
X_2
fixed domain
X_1

FIG. 1.4

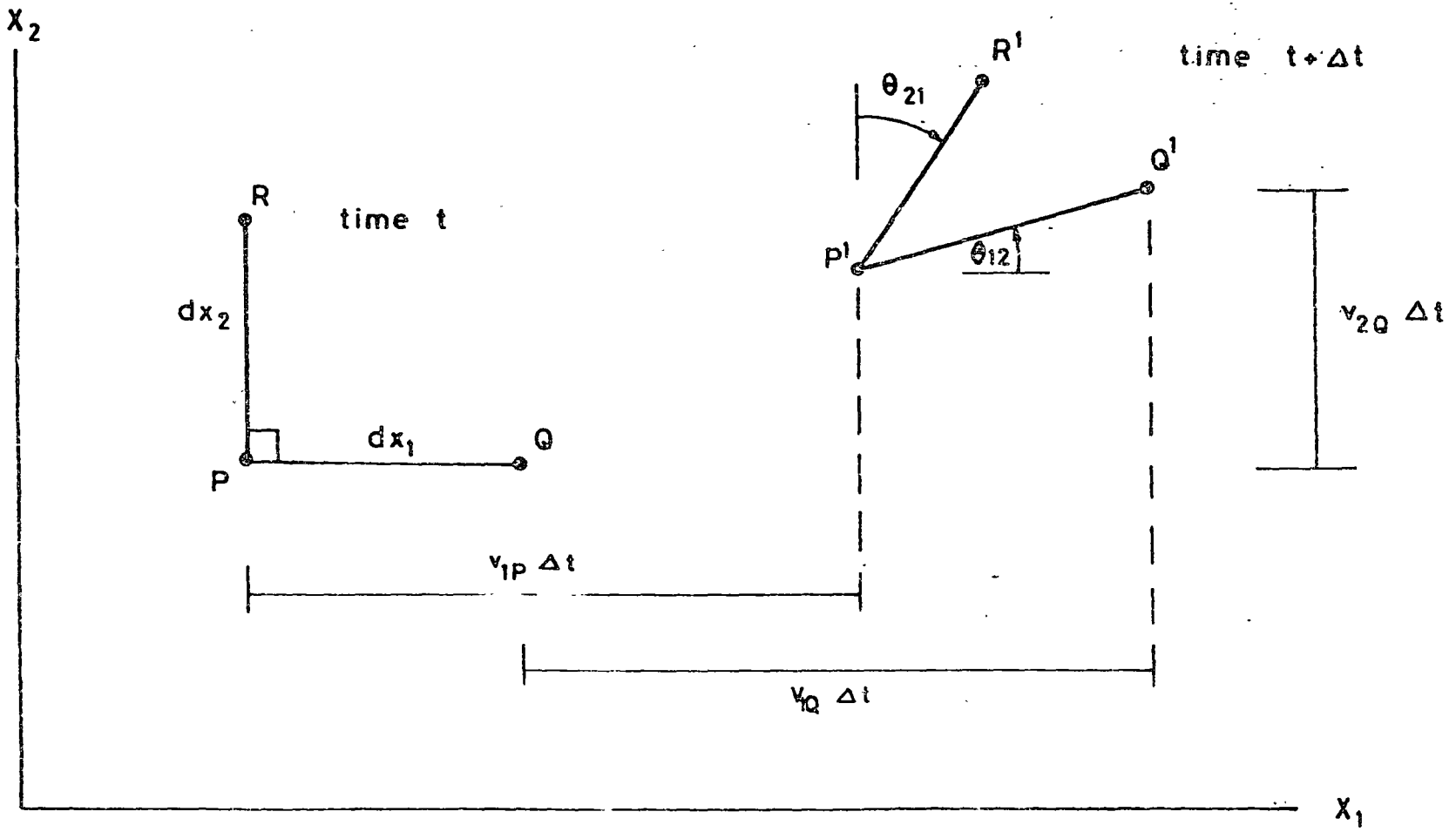

X2
X1
R
time t
dx2
dx1
Q
P
R1
time t + Δt
θ21
Q1
P1
θ12
v2Q Δt
v1P Δt
v1Q Δt

FIG.1.5

Governing Equations

CONTENTS

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

USOS DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPRESION EN AGUAS, LAGOS Y RIOS

INDICE DE ARTICULOS

MAYO, 1978.

JOURNAL OF THE HYDRAULICS DIVISION

FINITE ELEMENT SHALLOW LAKE CIRCULATION ANALYSIS

By Richard H. Gallagher,[1] F. ASCE, James A. Liggett,[2] M. ASCE, and Stevens T. K. Chan,[3] A. M. ASCE

INTRODUCTION

The finite element method has drawn increasing attention as a numerical analysis tool for fluid flow problems. The reasons for this growth of interest include the following: (1) Irregular boundaries can be treated accurately without computational difficulties or changes in programming or formulation of the method; (2) practical use can be made of widely available, general-purpose, finite element analysis programs which are virtually unlimited in the size of problem they can handle; and (3) known spacewise variations of physical properties can easily be taken into account.

Because of these advantages the finite element method is especially attractive as a practical method of analysis of lake circulation problems. In lake analysis irregular boundaries must be considered. The phenomenon described, is so complex in form that any numerical analysis procedure will entail hundreds, or perhaps thousands of unknowns. Due to the ease of transference of finite element programs from one computer facility to another, the development of relationships for this class of problem contributes to a capability which eventually may be applied routinely by analysts in many different and widely separated organizations.

As noted previously, variable physical properties are easily handled by the finite element method if their spatial distribution is known a priori. Thus, for

Note.--Discussion open until December 1, 1973. To extend the closing date one month, a written request must be filed with the Editor of Technical Publications, ASCE. This paper is part of the copyrighted Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, Vol. 99, No. HY7, July, 1973. Manuscript was submitted for review for possible publication on September 13, 1972.

[1]Prof. of Civ. and Environmental Engrg., Cornell Univ., Ithaca, N.Y.
[2]Prof. of Civ. and Environmental Engrg., Cornell Univ., Ithaca, N.Y.
[3]Research Assoc., Sch. of Civ. and Environmental Engrg., Cornell Univ., Ithaca, N.Y.

the subject problem, variations in eddy viscosity or the Coriolis acceleration can be taken into account if known. Variations in density present a more basic difficulty because the density distribution in a stratified lake (either through salt or temperature) is itself a part of the solution and cannot be specified a priori. The extension of the finite element method to account for such coupled phenomena is feasible, however, and the present work, on one side of the uncoupled problem is a step in this direction.

This paper presents a finite element formulation and numerical results for the analysis of the wind-induced steady-state circulation of variable-depth shallow homogeneous lakes. Formulative efforts and numerical results for finite element representations of lake and shallow basin circulation analysis have previously been described by Cheng (2) and by Loziak, Anderson, and Belytschko (10). These developments are exclusively two-dimensional, i.e., no account is taken of the variation of lake depth and the resulting velocities do not change with depth. Leonard and Melfi (6) present the theoretical relationships for a three-dimensional analysis which accounts for the velocity of the lake normal to the free surface, but no results are presented.

The present paper depends for its theoretical basis on a formulation of the governing differential equation that has been derived in detail by Liggett and Hadjitheodorou in Ref. 8. This development assumed homogeniety, hydrostatic pressure, specified wind shears, and small Rossby number. The latter assumption, together with a boundary condition of zero velocity normal to the lake free surface and the bottom, enables construction of a linear equation in two dimensions whose coefficients are a function of all three dimensions. Thus, the equation accounts for variable depth of the lake and for depthwise variation of velocity through numerical integration of equation coefficients that are functions of planform location.

The conventional basis for construction of a finite element representation is an integral form which, in the sense of a variational principle, corresponds to the governing differential equation. The transformation of the governing differential equation to integral form is accomplished here by use of the method of weighted residuals (3) rather than through variational calculus. The specific finite element representation employed is of triangular planform shape with an assumed linear variation of the stream function.

It should be noted that numerical solutions of the aforementioned governing differential equation, or of specialized forms of it, have previously been obtained with use of finite differences. Rectangular basins were analyzed in this manner in Refs. 5, 7, and 8 while Liu and Perez (9) solved the rectangular basin problem with removal of the Coriolis effect, i.e., with restriction to very shallow basins. The finite difference solutions are drawn upon herein to furnish comparison data for finite element solutions. In order to demonstrate the advantages of geometric representation, alluded to earlier, the finite element method is also applied herein to an analysis of the wind-driven circulation of Lake Ontario, for which no comparison results are available.

GOVERNING EQUATIONS

The purpose in this section is to present the governing differential equation for the steady-state, wind-driven circulation of shallow, homogeneous lakes,

as formulated by Liggett and Hadjitheodorou (8). Because detailed development of this equation is beyond the scope of this paper, interested readers should consult Ref. 8.

A cross section of the type of lake under study is pictured in Fig. 1. The origin of coordinates is fixed at the surface of the lake with z measured upwards.

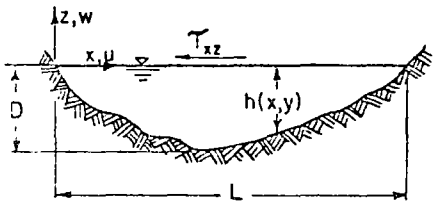


FIG. 1.—Representative Lake Cross Section

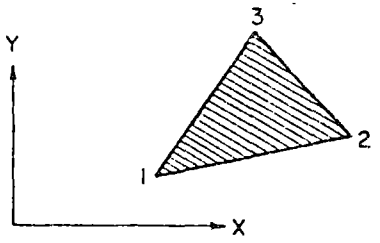


FIG. 2.Triangular Element

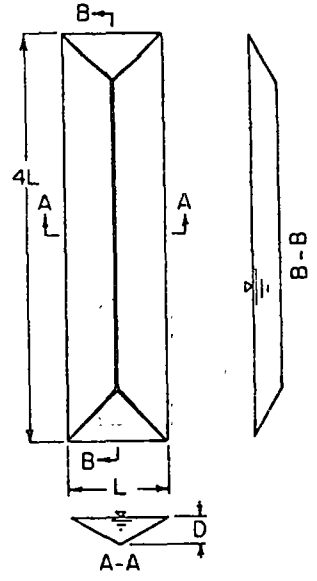


FIG. 3.—Rectangular Lake

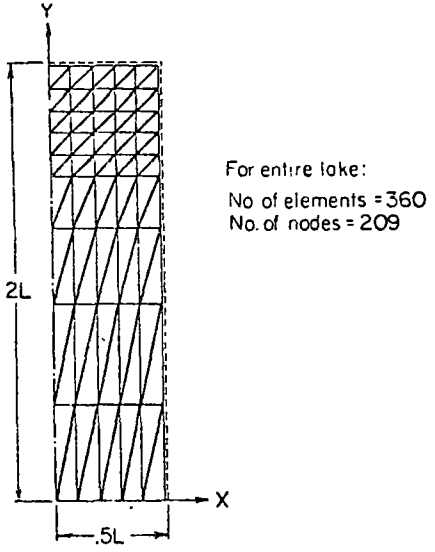


FIG. 4.—Finite Element Representation in Quadrant of Rectangular Lake

In accordance with the assumption of shallowness, i.e., hydrostatic pressure distribution, $D << L$. The eddy viscosity, η, and Coriolis parameter f are assumed constant in the formulation of the differential relationships. The distribution of pressure is assumed to be hydrostatic and surface wind stresses (τ_{xz}, τ_{yz}) are described. In order to linearize the problem the Rossby number (ratio of inertial forces to rotational forces) is taken to be small. The depths

to be used in a calculation are taken to be the actual depths under the assumed wind stress or, alternately, that the equilibrium depths are a sufficient approximation to the actual depth under the assumed wind stress. The x, y plane coincides with the water surface and $w \equiv 0$ at $z = 0$.

Many of the assumptions or approximations mentioned in connection with the present study have been evaluated by Liggett (7). Steady flow is assumed as the primary function of this paper is to indicate the utility of the finite element representation. This could be extended to an unsteady formulation in much the same way as was done in Ref. 7. The time response of a homogeneous lake, the effects of variable eddy viscosity, and the linearization were examined in Ref. 7. Unlike the unsteady problem, no rigid lid on the free surface is necessary if the depths are taken as the actual depths under the assumed wind stress. Even if the depths are taken as equilibrium depths, wind set-up can be computed from the resulting pressures.

Under the foregoing assumptions, the x, y, and z, direction momentum equations are of the form

$$-fv = -\frac{1}{\rho}\frac{\partial p}{\partial x} + \frac{\partial^2 u}{\partial z^2} \quad \ldots\ldots (1)$$

$$fu = -\frac{1}{\rho}\frac{\partial p}{\partial y} + \eta\frac{\partial^2 v}{\partial z^2} \quad \ldots\ldots (2)$$

$$g = -\frac{1}{\rho}\frac{\partial p}{\partial z} \quad \ldots\ldots (3)$$

in which u and v = the x and y direction velocities; ρ = the mass density per unit volume; and g = the acceleration due to gravity. The continuity equation is

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad \ldots\ldots (4)$$

and the boundary conditions relating to shear on the lake surface

$$\eta\frac{\partial u}{\partial z} = \tau_{xz}; \eta\frac{\partial v}{\partial z} = \tau_{yz} \quad \ldots\ldots (5)$$

and of zero velocity ($u = v = w = 0$) on all solid surfaces.

Operations on the preceding to produce a governing differential equation proceed as follows. First, the equations are written in nondimensional form through the introduction of an appropriate set of new variables. Then, a stream function, ψ, which satisfies the vertically integrated continuity condition corresponding to Eq. 4 is introduced. The stream function is defined as

$$\bar{u} = \frac{1}{h}\frac{\partial \psi}{\partial y}; \bar{v} = -\frac{1}{h}\frac{\partial \psi}{\partial x} \quad \ldots\ldots (6)$$

in which $\bar{u}$ and $\bar{v}$ = depthwise averages of the component velocities. Finally, the first three equations, with associated boundary conditions considered, are solved in terms of the stream function. The result is

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + A(x,y)\frac{\partial\psi}{\partial x} + B(x,y)\frac{\partial\psi}{\partial y} + C(x,y) = 0 \quad \ldots\ldots\ldots\ldots (7)$$

with the boundary condition that ψ is constant on the shore line. The coefficients, A, B, and C, in the equation are functions of the planform location (more specifically, functions of the lake bottom topography) as defined in Ref. 8, and C depends on the wind shear stresses as well.

The condition that ψ is constant on the lake boundary prescribes zero average velocity normal to the boundary. However, a stronger condition is necessary to insure that the point velocities normal to the boundary are everywhere zero. Such a condition is avoided by forbidding vertical boundaries. Thus the lake is confined by the surface and the bottom, on which all point velocities are specified as zero.

Finite Element Representation

There are three aspects to the establishment of the finite element equations: (1) Construction of integral relationships which correspond to the governing equations of the problem; (2) definition of the geometric form of the elements; and (3) representation of the assumed modes of behavior of the element.

In the finite element analysis of many physical problems, notably structural analysis, the preceding integral relationship is the stationary value of the functional, defining the variational (or energy) statement of the problem. The governing differential equations of the problem in terms of the independent variables of the functional are Euler equations of the functional. For certain circumstances the solution which yields a minimum value of the functional corresponds to the exact solution of the governing differential equation.

When the governing differential equations are not self-adjoint, as in the present case, there is doubt that a true variational statement of the problem can be constructed (1). Note further that variational principles do not have a pre-eminent, well-established position in fluid mechanics as they do in the approximate solution procedures in structural mechanics. The desired integral format for the subject problem is therefore established through application of the method of weighted residuals (4), noting that a particular form of this method gives identically the same integral relationship for problems which are self-adjoint.

The weighted residual concept assumes that an approximate representation of the independent variable, which in general does not satisfy the governing differential equation, will be chosen. In the present case this approximating trial function, $\bar{\psi}$, is of the form

$$\bar{\psi} = \sum_{i=1}^{n} N_i\psi_i = [N]\{\psi\} \quad \ldots\ldots\ldots\ldots (8)$$

in which ψ_i = a particular value of the independent variable and generally refers to such a value at the point i, and n = the chosen number of undetermined parameters ψ_i.

Designating the governing differential Eq. 7 as $L(\psi) = 0$, note that due to the approximate nature of $\bar{\psi}$ the result is

$$L(\bar{\psi}) = R \neq 0 \quad \ldots\ldots\ldots\ldots (9)$$

in which R = a residual value. Because the governing differential equation cannot be satisfied pointwise throughout the domain, V, of the problem its satisfaction is sought in the sense of a weighted average over the domain, i.e.

$$\int_V L(\bar{\psi})\,\phi\,dV = 0 \quad \ldots\ldots\ldots\ldots (10)$$

in which ϕ = the weighting function.

The weighting function may be specified in one of any number of forms. Here, the Galerkin form is chosen, in which the coefficients, N_i, of the trial function are employed. Each distinct trial function leads to a separate algebraic equation, using the procedure detailed in the following sections.

In the present case, simple triangular elements are used to represent the planform of a lake under consideration. The stream function is assumed to vary linearly in each element (Fig. 2), so that for this case Eq. 8 is of the form

$$\bar{\psi} = N_1\psi_1 + N_2\psi_2 + N_3\psi_3 \quad \ldots\ldots\ldots\ldots (11)$$

in which ψ_1, ψ_2, ψ_3 = the values of the stream function at the vertices; and N_1, N_2, N_3 = the corresponding shape functions. These functions are defined as

$$N_i = \frac{1}{2\Delta}(a_i + b_i x + c_i y) \quad \ldots\ldots\ldots\ldots (12)$$

with Δ = area of the triangular element ijk

$$a_i = x_j y_k - x_k y_j,\ b_i = y_j - y_k,\ c_i = x_k - x_j \quad \ldots\ldots\ldots\ldots (13)$$

in which i, j, k take the values of 1, 2, 3 cyclicly.

Applying Galerkin's criterion, the result is

$$\iint_\Delta \{N\}\left[\left(\frac{\partial^2[N]}{\partial x^2} + \frac{\partial^2[N]}{\partial y^2} + A\frac{\partial[N]}{\partial x} + B\frac{\partial[N]}{\partial y}\right)\{\psi\} + C\right]dxdy = 0 \quad \ldots\ldots\ldots\ldots (14)$$

Next, integration by parts is applied in the plane (Green's theorem). This operation reduces the order of the derivatives appearing in the integral and introduces the boundary terms into the resulting integral. In the present case the result is

$$\iint_\Delta \left[\left(-\frac{\partial\{N\}}{\partial x}\frac{\partial[N]}{\partial x} - \frac{\partial\{N\}}{\partial y}\frac{\partial[N]}{\partial y} + A\{N\}\frac{\partial[N]}{\partial x} + B\{N\}\frac{\partial[N]}{\partial y}\right)\{\psi\} + \{N\}C\right]dxdy + \oint\{N\}\frac{\partial[N]}{\partial n}\{\psi\}\,dS = 0 \quad \ldots (15)$$

The values ψ_i = zero on the entire exterior boundary in the present problem and the closure integrals along internal (interelement) boundaries vanish as ele-

ment size decreases (13), or if continuity of $\partial\{N\}/\partial n$ across element boundaries is preserved. Thus, the contour integral term is excluded from subsequent consideration. Evaluation of Eq. 15 then yields the system of equations

$$[k^e]\{\psi\} = \{r^e\} \quad \ldots\ldots (16)$$

in which

$$[k^e] = \iint_\Delta \left(-\frac{\partial\{N\}}{\partial x}\frac{\partial\{N\}}{\partial x} - \frac{\partial\{N\}}{\partial y}\frac{\partial\{N\}}{\partial y} + A\{N\}\frac{\partial\{N\}}{\partial x} + B\{N\}\frac{\partial\{N\}}{\partial y}\right) dxdy \quad \ldots\ldots (17)$$

$$\{r^e\} = -\iint_\Delta \{N\}\, C dxdy \quad \ldots\ldots (18)$$

Note should be taken of certain aspects of the numerical evaluation of Eq. 17. First, due to the terms $A\{N\}(\partial\{N\}/\partial x)$ and $B\{N\}(\partial\{N\}/\partial y)$, the resulting algebraic equations will be nonsymmetric. This means that advantage can not be taken of symmetricity as encountered in most structural finite element analysis. That is, every term of the element matrix has to be evaluated and the *entire* banded system matrix has to be stored for computing the solution.

As Eq. 7 indicates, the coefficients A, B, and C are functions of x and y. Herein the decision is made to approximate the variation of these terms within each element by linear functions, similar to Eq. 11. Choice of a constant value for each element, say at the centroid, would simplify integration of Eqs. 17 and 18. Proper comparison of the finite difference solutions of Ref. 8, in which these coefficients vary between the points of the mesh, is being sought, however. It should further be observed that integration within the triangle is simplified considerably by use of area coordinates (12).

The equations of the complete lake are constructed from the equations of the elements by imposing the condition of stream function continuity at each element joint, which is synonymous with simple addition of all coefficients (k_{ij} and r_i) with like subscripts. Thus, the full set of equations is of the form

$$[K]\{\psi\} = \{R\} \quad \ldots\ldots (19)$$

in which

$$K_{ij} = \Sigma k_{ij} \quad \ldots\ldots (20)$$

$$R_i = \Sigma r_i \quad \ldots\ldots (21)$$

and the summations range over all elements with terms with the subscripts i and j.

After solution of Eq. 19 for $\{\psi\}$, other variables, such as averaged velocities, pressure gradients, and velocities at different depth levels can be subsequently evaluated by back substitutions. Herein, because a linear field in ψ has been assumed, its derivatives $\partial\psi/\partial x$, $\partial\psi/\partial y$ are constants in each element. Thus, from the definition of ψ, the average velocities $\bar{u}$ and $\bar{v}$ are constant within each element. From Ref. 8 the point velocities as a function of planform and depth are

$$u = -\frac{\partial p}{\partial y} \quad \text{os } mz(c_2 e^{mz} - c_4 e^{-mz}) - \sin mz(c_1 e^{mz} - c_3 e^{-mz}) \quad \ldots\ldots (22a)$$

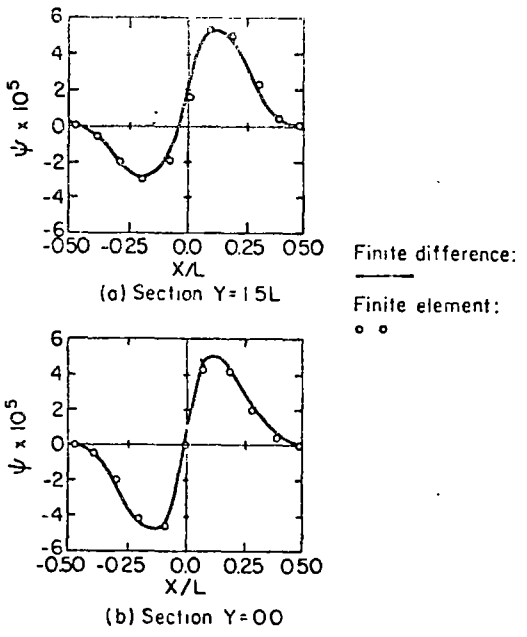


FIG. 5.—Comparison of Stream Function Solutions for Rectangular Lake at Representative Sections

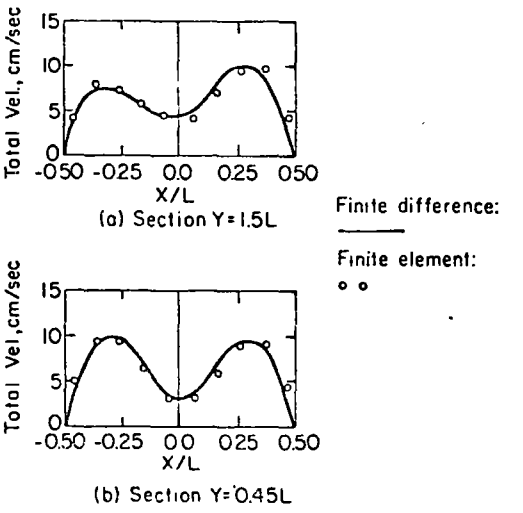


FIG. 6.—Comparison of Total Velocity Solutions on Rectangular Lake Surface at Representative Sections

$$v = \frac{\partial p}{\partial x} + \cos mz(c_1 e^{mz} + c_3 e^{-mz}) + \sin mz(c_2 e^{mz} + c_4 e^{-mz}) \quad \ldots\ldots (22b)$$

in which the terms m, c_1, c_2, c_3, and c_4 = functions of x and y. The reader is referred to Ref. 8 for the exact definitions of these terms, which are rather complex. The same paper expresses the pressure gradients as a function of x and y and the derivatives of ψ.

Numerical Results

Two problems are solved as an illustration of the present approach. The first problem, shown in Fig. 3, enables comparison with finite difference results (8). This idealized lake is oriented in a north-south direction with a length four times the width. The following values were employed in numerical calculation: $f = 0.0001$ rad/s; $D = 8{,}000$ cm; $\eta = 200$ cm^2/s; $L = 1.25 \times 10^7$ cm; $\tau = 1.0$ cm^2/s^2; and $g = 980$ cm/s^2.

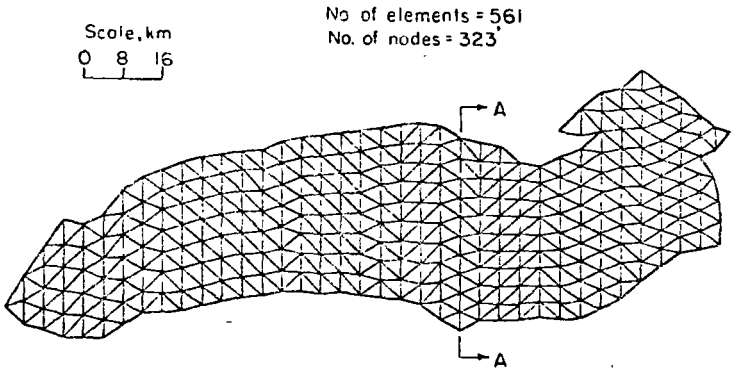


FIG. 7.—Finite Element Representation of Lake Ontario

The finite element representation of a quadrant of this lake is shown in Fig. 4. A total of 90 elements are arrayed in the quadrant but four times this number, 360, with 209 joints, and therefore the same number of equations, were employed in actual computation because the geometric symmetry about the x and y-axes does not apply to the circulation behavior being calculated.

As is apparent from the definition of ψ, a zero depth represents a computational singularity and it is necessary to have a finite, but small, depth all along the boundary. The flow region under analysis was, therefore, taken to be one bounded by a contour of 5% maximum depth, a value which has been found to be adequate in previous numerical solutions, and the flow exterior to the boundary is ignored (assumed to be at rest).

Figs. 5 and 6 show results for ψ and the velocity resultant at selected cross sections for the case of a south wind acting on the lake surface. Also shown are the finite difference results from Ref. 8, in which 1,701 equally-spaced pivotal points were used. The finite difference and finite element results are seen to be in close agreement. It should be noted that these comparisons are presented to affirm the validity of the finite element method in solution of

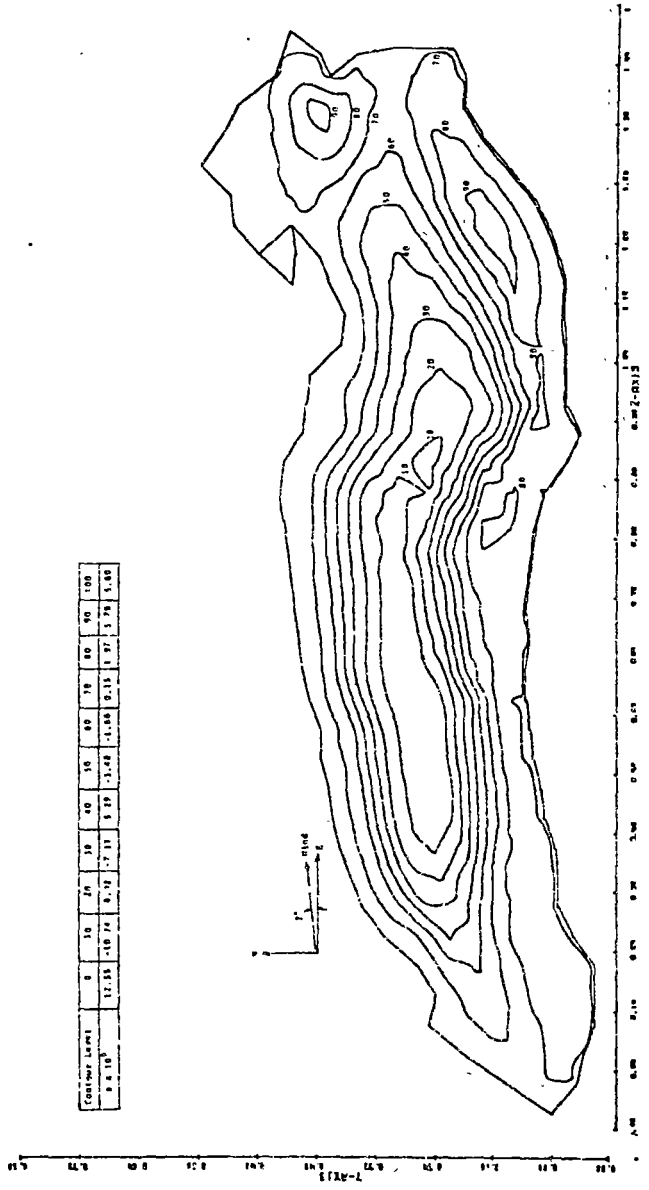


FIG. 8.—Lake Ontario—Contour Plot of Stream Function

this problem and not to measure the relative efficiency of the finite element and finite difference procedures. A bare comparison of computational effort, based on the number of equations to be solved would not be realistic due to such factors as the relative effort in forming the equations and the narrow bandwidth of finite difference equations.

Also, it is quite possible that finite difference results, which would prove comparable with the finite element solution, could have been obtained with much fewer than 1,701 pivotal points. The computational costs of the two solutions cannot be compared due to significant differences in the computer hardware and software employed for the respective cases.

The second problem for which numerical results are described herein is the calculation of the circulation of Lake Ontario due to a wind shear prevailing in the local average direction at Rochester in February, as shown in Fig. 8.

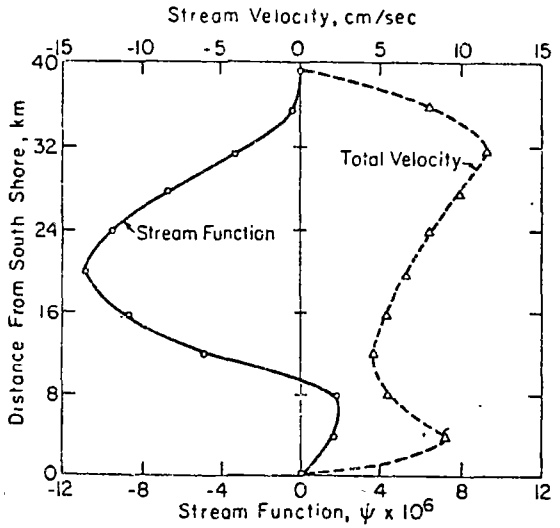


FIG. 9.—Stream Function and Total Velocity Solution on Lake Ontario Surface at Section A-A

A careful representation of the geometry and bottom topography of this lake was compiled by Canada Center for Inland Waters (2). These data are employed herein to define a finite element representation consisting of 561 triangles joined at 323 points. The specified wind shear stress and the physical constants f, η, and g are the same as in the first problem.

Fig. 8 shows contours of the stream function in the circulating lake. This figure has been generated by a contour plotting routine which is part of the computer program. A plot of the distribution of the stream function and velocity on a representative north-south section across the lake appears in Fig. 9.

No comparison results are available for this problem. Although the correct Coriolis parameter was used, no attempt was made to choose a physically accurate eddy viscosity or to represent ice formation or variation of wind stress. It is unlikely that field measurements of the form necessary for comparison

purposes will be available in the future. Large-scale modeling is a promising alternative source of comparison data but no such data exist yet for this lake and when they are obtained it is to be expected that limitations on representation of the pertinent dimensionless ratios [see Rumer and Hoopes (11)] will require somewhat different conditions on the comparison analyses than those employed herein.

In computational aspects, execution times for the finite element solutions of the preceding problems were generally small, between 1/2 min and 1-1/2 min on an IBM 360/65.

One method of establishing confidence in the validity of the present results is by performance of further analyses with either a revised gridwork or with higher-order elements on the same gridwork. Work in this direction is in progress. Correspondence of the results of these solutions will add to confidence in the accuracy of the solution of Eq. 7 for this situation but cannot, of course, demonstrate that this differential equation properly describes the behavior of the actual lake.

Summary and Conclusions

The finite element method has been shown to be effective in the analysis of lake circulation. Such problems are quite complicated from a geometric standpoint and a realistic analysis with use of any method must inevitably require a large-scale computation. The finite element method is attractive in this respect because of the possibility of using existing large-scale, general-purpose, finite element computer programs. It is especially promising as the basis for analysis of more complex circulation phenomena in lakes, such as the response to the introduction of a thermal plume into a stratified lake. Extension of the present work to three dimensions in a more general way, with removal of the assumptions which produced a two-dimensional differential equation, can be accomplished without extension of basic theory. Appropriate trial functions for such elements are reviewed in Ref. 4. The isoparametric element concept (14), in which higher-order polynomial trial functions are also employed to map curvilinear element boundaries, is especially attractive as a means of improving the efficiency of three-dimensional elements in geometric representation of a problem. The computational expense of three-dimensional representations is inevitably vastly increased in comparison with two-dimensional models, however, their application must necessarily be motivated by a desire to include new phenomena.

Acknowledgment

Work described herein was supported by the National Science Foundation under Research Grant GK-23992.

Appendix I.—References

1. Aral, M., Mayer, P. G., and Smith, C. V., "Finite Element Galerkin Method Solutions to Selected Elliptic and Parabolic Differential Equations," *Proceedings of 3rd Conference on Matrix Methods in Structural Mechanics*, Dayton, Ohio, 1971.
2. Cheng, R. T., "Numerical Investigation of Lake Circulation Around Isla by the

Finite Element Method," *International Journal for Numericals Methods in Engineering*, Vol. 5, No. 1, 1972.
3. Finlayson, B., *The Method of Weighted Residuals and Variational Principles*, Academic Press, New York, N.Y., 1972.
4. Gallagher, R. H., "Computational Methods in Nuclear Reactor Structural Design for High-Temperature Applications—An Interpretive Report," Chap. 3, *Report Oak Ridge National Lab., AEC-4756*, July, 1972.
5. Lee, K. K., and Liggett, J. A., "Computation of Wind-Driven Circulation in Stratified Lakes," *Journal of the Hydraulics Division*, ASCE, Vol. 96, No. HY10, Proc. Paper 7634, Oct., 1970, pp. 2089-2115.
6. Leonard, J., and Melfi, D., "3-D Finite Element Model for Lake Circulation," *Proceedings of the 3rd Conference on Matrix Methods in Structural Mechanics*, Dayton, Ohio, 1971.
7. Liggett, J. A., "Cell Method for Computing Lake Circulation," *Journal of the Hydraulics Division*, ASCE, Vol. 96, No. HY3, Proc. Paper 7152, Mar., 1970, pp. 725-743.
8. Liggett, J. A., and Hadjitheodorou, C., "Circulation in Shallow Homogeneous Lakes," *Journal of the Hydraulics Division*, ASCE, Vol. 95, Proc. Paper 6454, Mar., 1969, pp. 609-620.
9. Liu, H., and Perez, H., "Wind-Induced Circulation in Shallow Water," *Journal of the Hydraulic Division*, ASCE, Vol. 97, No. HY7, Proc. Paper 8235, July, 1971, pp. 923-935.
10. Loziuk, L. A., Anderson, J. C., and Belytschko, T., "Finite Element Approach to Transient Hydrothermal Analysis of Small Lakes," presented at the October 16-22, 1972, ASCE Annual and National Environmental Engineering Meeting, held at Houston, Tex. (Preprint 1799).
11. Rumer, R., and Hoopes, J. A., "Modeling Great Lakes Circulations," *Report*, Wisconsin Water Resources Center, 1971.
12. Silvester, P., "High-Order Polynomial Triangular Elements for Potential Problems," *International Journal of Engineering Science*, Vol. 7, 1969, pp. 849-861.
13. Szabo, B., and Lee, G. C., "Derivation of Stiffness Matrices for Problems in Plane Elasticity by Galerkin's Method," *International Journal for Numerical Methods in Engineering*, Vol. 1, 1969, pp. 301-310.
14. Zienkiewicz, O. C., "Isoparametric and Allied Numerically Integrated Elements—A Review," *Proceedings of the Office of Naval Research International Symposium on Numerical and Computer Methods in Structural Mechanics*, A. Robinson, ed., Academic Press, New York, N.Y., 1972.

Appendix II.—Notation

The following symbols are used in this paper:

A, B, C = coefficients in Eq. 7 as defined in Ref. 8;
a_i, b_i, c_i = quantities relating coordinates of joints in element;
c_1, c_2, c_3, c_4 = coefficients for evaluating velocity components as defined in Ref. 8;
D = typical vertical dimension used to normalize depth;
f = Coriolis parameter;
g = acceleration of gravity;
h = normalized depth of lake;
$[K]$ = coefficient matrix of resulting system equations;
$[k^e]$ = 3×3 element matrix as defined by Eq. 17;
L = typical horizontal dimension used to normalize horizontal dimensions;
$L(\psi)$ = linear operator to operate on ψ;
m = $Df/2\eta$;

N_i = shape function at joint i;
n = chosen number of undetermined parameters ψ_i;
p = local pressure;
R = right-hand side of system equations or residual;
$\{r^e\}$ = element column matrix as defined by Eq. 18;
u, v, w = velocity components in x, y, and z directions, respectively;
$\bar{u}, \bar{v}$ = average velocity components in x and y directions, respectively;
V = entire flow domain under consideration;
x, y, z = Cartesian coordinates with x positive eastward, y positive northward, and z positive upward and zero at surface;
Δ = area of triangular element;
η = eddy viscosity;
ρ = fluid density;
τ_{xz}, τ_{yz} = surface wind stresses in x and y directions, respectively;
ψ = stream function; and
$\bar{\psi}$ = approximate stream function solution.

JOURNAL OF THE ENGINEERING MECHANICS DIVISION

CONVECTIVE TRANSPORT FINITE ELEMENT ANALOG

By Keith W. Bedford,[1] A. M. ASCE and James A. Liggett,[2] M. ASCE

INTRODUCTION

The simultaneous prediction of momentum, heat, or mass transfer in closed cavities has challenged researchers for years Certain flow situations marked by extreme diffusion or high inertia can be mathematically analyzed by closed-form solution or a boundary layer analysis (5,17,31). When a flow problem is not dominated by either physical process, mathematical decoupling of the momentum transport from the heat (or density) transport is prohibited. The resulting natural or combined convection cavity problem (17), difficult as it is, quite often is further complicated by the occurrence of circulation cells accompanied by high shear rates and density gradients (4,8,29,34). Since the governing simultaneous equations are nonlinear they require careful numerical analysis.

The finite element method will be used to form a numerical analog for the viscous cavity problem. A review of finite difference schemes for convection problems is found in Ref. 25. Several features of the physical problem are particularly suited for analysis by the finite element method (FEM) such as: (1) Irregular boundaries are treated accurately without computational difficulty; (2) variable boundary conditions such as differential heat input, temperature distribution, or wind shear are easily handled by the FEM; and (3) the solution of nonlinear fluid problems by FEM is just beginning (4,21,22,23,32), but already excellent iterative stability and rapid convergence are apparent.

The first fluid mechanics application of the FEM was to linear potential flow problems (2,33). Extensions to lubrication and creep flow followed thereafter (1,9,10). Solutions of viscous homogeneous flow problems are now available (6,7,15,16,20,27). Nonlinear viscous flow problems, being the most difficult, are only beginning to receive attention. Olson (29) presents a quintic element

Note.—Discussion open until May 1, 1976. To extend the closing date one month, a written request must be filed with the Editor of Technical Publications, ASCE. This paper is part of the copyrighted Journal of the Engineering Mechanics Division, Proceedings of the American Society of Civil Engineers, Vol. 101, No. EM6, December, 1975. Manuscript was submitted for review for possible publication on August 2, 1974.

[1]Asst. Prof. of Civ. Engrg., The Ohio State Univ., Columbus, Ohio.
[2]Prof. of Civ. and Environmental Engrg., Cornell Univ., Ithaca, N.Y.

Newton-Raphson FEM procedure to analyze several homogeneous flows for the Reynolds number as high as 1,000. Stable iteration and rapid convergence are particularly noticeable in this method. Skiba, Unny, and Weaver (32) present a rectangular element steady-state solution for natural convection in a slot. A weighted average method was used to iterate velocity and temperature values, causing iteration and convergence behavior to be inferior to Olson's method. King, Norton, and Orlob (23) present a Newton-Raphson technique to analyze stratified flow over a broad-crested weir.

In this work a finite element analog is presented for the class of steady, viscous, incompressible, two-dimensional, heat, or mass transfer flow problem. The Galerkin method of weighted residuals (MWR) is used to derive a functional. A cubic plate bending element is used with the streamfunction, ψ, and temperature or density (T or ρ) as unknowns. The element provides nodal continuity for not only ψ but more importantly the flow velocities which are derivatives of ψ. An amalgamation of Olson's (29) Newton-Raphson method and Skiba's (32) weighted average techniques is used to iterate the coupled system of equation. The results from three cases are presented. These include homogeneous linear shear driven cavity flow, lateral temperature gradient induced natural convection in a box, and shear driven stably stratified cavity flow.

CONSERVATION EQUATIONS

The three problems analyzed in this paper are treated by the same equations. The coordinate system is indicated in Fig. 1.

Only steady-state problems are considered. Laminar friction is used although the viscosity maybe interpreted as an eddy viscosity which is held constant in the spirit of mathematical tractibility. Density variations are assumed negligible except in the buoyancy terms (the Boussinesq approximation), and the density is unaffected by pressure (incompressible fluid).

General Form of Basic Equations.—The general form of the two dimensional steady-state shear or buoyancy driven cavity flow equations is as follows. The continuity equation is

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 \quad \ldots\ldots (1)$$

The x-momentum equation is

$$u\frac{\partial u}{\partial x} + w\frac{\partial u}{\partial z} = -\frac{1}{\rho_o}\frac{\partial p}{\partial x} + \eta\nabla^2 u \quad \ldots\ldots (2)$$

The z-momentum equation is

$$u\frac{\partial w}{\partial x} + w\frac{\partial w}{\partial z} = -\frac{1}{\rho_o}\frac{\partial p}{\partial z} + \eta\nabla^2 w - \frac{\rho}{\rho_o}g \quad \ldots\ldots (3)$$

The diffusion-convection equation is

$$\frac{\partial \phi}{\partial x} + w\frac{\partial \phi}{\partial z} = \alpha\nabla^2\phi \quad \ldots\ldots (4)$$

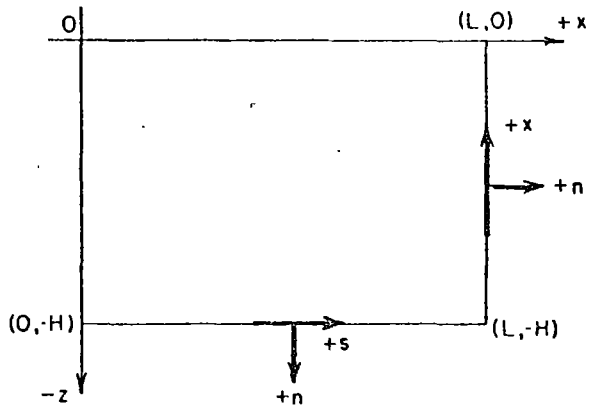


FIG. 1.—Notation and Coordinate Systems for Rectangular Cavities

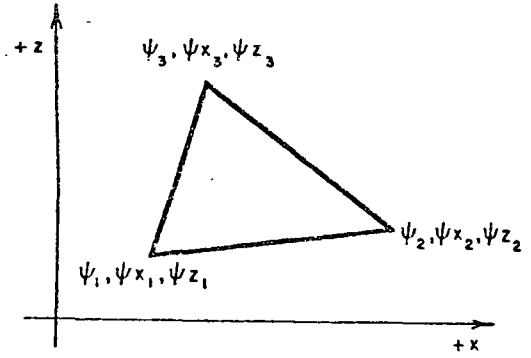


FIG. 2.—Element Degrees-of-Freedom

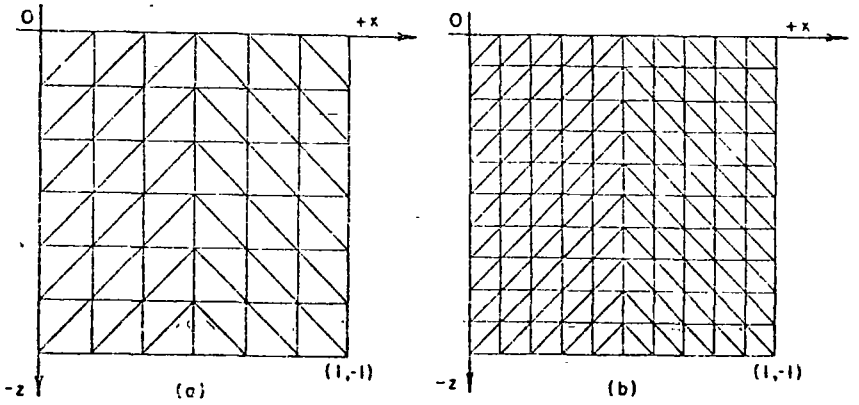


FIG. 3.—Cavity Discretizations

in which u and w are the velocities in the x and z directions, respectively; P = the pressure; ρ = the density; ϕ is either the temperature, T, or the mean turbulent density; and η, α, and ρ_o are the viscosity, diffusivity, and reference density.

Governing Cavity Flow Equations.—The equations are written in nondimensional form by defining the following variables:

$$x^* = \frac{x}{H};\quad z^* = \frac{z}{H};\quad u^* = \frac{u}{U};\quad w^* = \frac{w}{U};\quad \eta^* = \frac{\eta}{\eta_o};\quad \alpha^* = \frac{\alpha}{\alpha_o};$$

$$\phi = \frac{-(\phi - \phi_o)}{\phi_T - \phi_B} = \frac{-(\phi - \phi_o)}{\Delta\phi};\quad p^* = \frac{p}{\frac{1}{2}\rho_o U^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

in which H, U, α_o, η_o, and ρ_o represent reference values; and ϕ_T and ϕ_B are conservation quantities at the cavity top and bottom. Upon substitution of these variables into the equations the asterisks are dropped, and all variables from this point are dimensionless (except in defining R, Pr, Gr).

Eqs. 2 and 3 are cross differentiated, thereby eliminating the pressure and forming a vorticity equation. Finally, a pair of coupled equations emerge which will be used for the solution:

$$-\frac{\eta}{R}\nabla^4\psi + \frac{\partial(\nabla^2\psi,\psi)}{\partial(z,x)} - \frac{Gr}{R^2}\frac{\partial\phi}{\partial x} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

$$\frac{-\alpha}{PrR}\nabla\phi^2 + \frac{\partial(\phi,\psi)}{\partial(z,x)} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

in which ψ = the streamfunction ($u = \partial\psi/\partial z$, $w = -\partial\psi/\partial x$); $Pr = \eta/\alpha$ (Prandtl number), $R = UH/\eta_o$ (Reynolds number); and $Gr = (\nabla\rho)gH^3/\rho_o\eta_o^2$ (Crashof number). The nonlinear terms represented by the Jacobian make the equations difficult to solve. The adequate solution of the problem requires full retention of these terms in the numerical method.

Numerical Technique

The development of the FEM analog begins with the derivation of a functional form of the governing equations by the Galerkin method of weighted residuals (MWR). The remainder of this section considers in order: (1) formulation of the extremum principle; (2) the element and the element stiffness matrices; (3) the iterative solution technique; (4) boundary conditions; and (5) other methods.

Formulation of Extremum Principle.—Despite derivations of nonlinear "variational principles" (11,13,14,18,19,24) a true variational formulation is unknown for this nonlinear problem. Although proofs of boundedness and convergence do not exist as in the linear case, the Galerkin MWR provides a direct formulation of an integral expression for the system of nonlinear equations. The concept assumes that an approximation to the dependent variables (ψ,ϕ) exists and is of the form $\bar{\psi} = \Sigma_{i=1}^{n} N_i\psi_i$; $\bar{\phi} = \Sigma_{i=1}^{n} N_i\phi_i$, in which ψ_i (or ϕ_i) is a particular value of the variable at point i, and N_i functions of x and z are called shape

functions. If the governing equation, Eq. 6, is designated $D_1(\bar{\psi},\bar{\phi})$ then the substitution of the series, Eqs. 3 and 9, gives $D_1(\bar{\psi},\bar{\phi}) = R_1^* \neq 0$, in which R_1^* is the residual caused by the approximate nature of the series. The Galerkin criterion specifies that the weighted average of the residual over the domain of the cavity be zero, thus $\int_A D_1(\bar{\psi},\bar{\phi})\, N_i\, dA = 0$, in which the weight function is chosen as the shape function, N_i.

Eqs. 6 and 7 are multiplied by the weight functions, N_i, and integrated over the cavity, the result after integration by parts is

$$\frac{-\eta}{R}\iint\left(\frac{\partial^2 N_i}{\partial z^2}\frac{\partial^2\psi}{\partial z^2}+2\frac{\partial^2 N_i}{\partial x\partial z}\frac{\partial^2\psi}{\partial x\partial z}+\frac{\partial^2 N_i}{\partial x^2}\frac{\partial^2\psi}{\partial x^2}\right)dzdx$$

$$+\iint\left(-\frac{\partial N_i}{\partial x}\frac{\partial\psi}{\partial z}+\frac{\partial N}{\partial z}\frac{\partial\psi}{\partial x}\right)\nabla^2\psi\, dxdz$$

$$-\frac{Gr}{R^2}\iint N_i\frac{\partial\phi}{\partial x}dxdz-\frac{\eta}{Re}\oint\frac{\partial N_i}{\partial z}\frac{\partial^2\psi}{\partial z^2}dx=0 \quad\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

and from Eq. 7

$$\frac{\alpha}{PrR}\iint\left(\frac{\partial N_i}{\partial x}\frac{\partial\phi}{\partial x}+\frac{\partial N_i}{\partial z}\frac{\partial\phi}{\partial z}\right)dxdz$$

$$+\iint\left(N_i\frac{\partial\psi}{\partial z}\frac{\partial\phi}{\partial x}-N_i\frac{\partial\psi}{\partial x}\frac{\partial\phi}{\partial z}\right)dxdz=0 \quad\ldots\ldots\ldots\ldots\ldots\ldots (9)$$

The boundary integrals arising from the integration by parts provide the device by which boundary conditions are imposed.

Element Formulation.—The element to be used in this paper is the nine degree-of-freedom nonconforming plate bending triangle presented by Bazeley, et al. (3). The element has the unknown ψ and its first derivatives as nodal unknowns; therefore, the streamfunction is immediately differentiable to obtain the velocities. Previous application of cubic elements to flow problems (6,15) also shows the element to be accurate and reduces the number of elements necessary to describe the system. This element is a slightly altered form of the 10 degree-of-freedom element (3,6,12,15) in that the centroid node is distributed among the nine corner nodes. For nonlinear iterative analysis the computational difficulties of the centroid nodes are severe and require its elimination. The element though nonconforming appears to give satisfactory first cut approximations for initial work on flow problems.

To establish the cubic polynomial the unknown ψ (or ϕ) and its derivatives, ψ_x and ψ_z, are defined at the three corner nodes (Fig. 2). If area coordinates are defined as $L_i = A_i/A$, in which A = total area of the element, and A_i = the area of the triangular subregion (i = 1,2,3), then the shape function description is as given by Ref. 3:

$$\left.\begin{aligned}\psi_{e_j} &= N_1\psi_1 + N_2\psi x_1 + N_3\psi z_1 + \ldots + N_9\psi z_3\\ \phi_{e_j} &= N_1\phi_1 + N_2\phi \quad + N_3\phi z_1 + \ldots + N_9\phi z_3\end{aligned}\right\} \quad\ldots\ldots\ldots\ldots (10)$$

in which $N_1 = L_1^3 + 3L_1^2(L_2+L_3) + 2L_1L_2L_3$;

$$N_2 = c_3\left(L_1^2L_2+\frac{1}{2}L_1L_2L_3\right)-c_2\left(L_3L_1^2+\frac{1}{2}L_1L_2L_3\right);$$

$$N_3 = -b_3\left(L_1^2L_2+\frac{1}{2}L_1L_2L_3\right)+b_2\left(L_3L_1^2+\frac{1}{2}L_1L_2L_3\right) \quad\ldots\ldots (11)$$

and $a_1 = x_2z_3 - x_3z_2$, $b_1 = z_2 - z_3$, $c_1 = x_3 - x_2$. N_4 and N_7 are found by cyclic permutation of indices in Eq. 11 similarly, N_5 and N_8 are found from N_2 and N_6 and N_9 are found from N_3.

First and second-order differentiations are now performed and the resulting matrices are substituted into the functional expressions, Eqs. 10 and 11, yielding two sets of nonlinear algebraic equations for the element, e (see Ref. 35). Thus

$$-\frac{\eta}{R}\Phi_{ij}\psi_j+\theta_{ikj}\psi_k\psi_j+\frac{Gr}{R^2}\lambda_{ij}\phi_j-P_i^e=0 \quad\ldots\ldots\ldots\ldots\ldots\ldots (12)$$

$$\frac{\alpha}{PrR}\xi_{ij}\phi_j+\pi_{ikj}\psi_k\phi_j-Q_i^e=0 \quad\ldots\ldots\ldots\ldots\ldots\ldots (13)$$

Column vectors P_i^e and Q_i^e are formulated from the boundary integrals in Eqs. 8 and 9. It is assumed that these vectors are zero for all element sides not on the physical boundary of the problem. The customary procedure is to assemble the element stiffness matrices (10,12) into a global systems of equations, treating the nonlinear matrices, θ and π, as quasilinear in the summation process. Judicious application of the iteration technique would then result in acceptable results. For reasons given in a later section, the solution process is reversed, i.e., the iteration or perturbation technique is employed at the element level and a global system is then composed.

Iterative Solution Technique.—The solution of Eqs. 12 and 13 is by a Newton-Raphson technique coupled with a weighted averaging method. This method is an amalgamation of Olson's (29) and Skiba's (32) work.

The method begins by applying the Newton-Raphson procedure at the element level. Let $\psi_{o_j}^n$ and $\phi_{o_j}^n$ be the nth approximation to the correct solution, ψ_j^e and ϕ_j^e (for element e) of Eqs. 12 and 13. If the ith equation in Eq. 12 is f_i and the ith equation in Eq. 13 is g_i then by a truncated Taylor series:

$$f_i(\psi_{o_j}^{n-1},\phi_{o_j}^{n-1})+\sum_{l=1}^{9}\left(\frac{\partial f_i}{\partial\psi_l}\right)\Delta\psi_l^n \simeq f_i(\psi_j^e,\phi_j^e)=0 \quad\ldots\ldots\ldots\ldots (14)$$

and

$$g_i(\psi_{o_j}^{n-1},\phi_{o_j}^{n-1})+\sum_{l=1}^{9}\left(\frac{\partial g_i}{\partial\phi_l}\right)\Delta\phi_l^n \simeq g_i(\psi_j^e,\phi_j^e)=0 \quad\ldots\ldots\ldots\ldots (15)$$

in which $$\frac{\partial f_i}{\partial\psi_j}=\frac{-\eta}{R}\phi_{ij}+\sum_{l=1}^{9}(\theta_{ikl}+\theta_{ijk})\,\psi_{ok}^{n-1} \quad\ldots\ldots\ldots\ldots (16)$$

$$\frac{\partial g_l}{\partial\phi_j} \quad \frac{\alpha}{3Pr}\xi_{ij}+\sum_{l=1}^{9}\pi_{ikj}\psi_{ok}^{n-1} \quad\ldots\ldots\ldots\ldots (17)$$

$$\Delta\psi_i^n = \psi_{oi}^n - \psi_{oi}^{n-1} \quad \ldots\ldots (18)$$

$$\Delta\phi_i^n = \phi_{oi}^n - \phi_{oi}^{n-1} \quad \ldots\ldots (19)$$

and

$$\left.\begin{aligned} f_i(\psi_{oj}^{n-1}, \phi_{oj}^{n-1}) &= \frac{-\eta}{R}\Phi_{ij}\psi_{oj}^{n-1} + \theta_{ikj}\psi_{ok}^{n-1}\psi_{oj}^{n-1} + \frac{Gr}{R^2}\lambda_{ij}\phi_{oj}^{n-1} \\ g_i(\psi_{oj}^{n-1}, \phi_{oj}^{n-1}) &= \frac{\alpha}{PrR}\xi_{ij}\phi_{oj}^{n-1} + \pi_{ikj}\psi_{ok}^{n-1}\phi_{oj}^{n-1} \end{aligned}\right\} \quad \ldots\ldots (20)$$

A final set of equations for the element e perturbation values is written

$$S_{ij}^e \Delta\psi_j^n = -f_i^{n-1}; \quad T_{ij}^e \Delta\rho_j^n = -g_i^{n-1} \quad \ldots\ldots (21)$$

in which $S_{ij}^e = \partial f_i / \partial \psi_j$; and $T_{ij}^e = \partial g_i / \partial \phi_j$.

The systems of element "slope" matrices can be assembled into a set of global "slope" matrices

$$S_{ij}^{n-1}\Delta\psi_j^n = -F_i^{n-1} \quad \ldots\ldots (22)$$

$$T_{ij}^{n-1}\Delta\phi_j^n = -G_i^{n-1} \quad \ldots\ldots (23)$$

in which, if there are p nodes in the system, $\Delta\psi_j^n$(and similarly $\Delta\phi_j^n$) equals

$$\Delta\psi_j^n = \Delta\psi_1, \frac{\partial\Delta\psi_1}{\partial x}, \frac{\partial\Delta\psi_1}{\partial z}, \ldots, \Delta\psi_p, \frac{\partial\Delta\psi_p}{\partial x}, \frac{\partial\Delta\psi_p}{\partial z} \quad \ldots\ldots (24)$$

A weighted average scheme completes the iteration process. When Eq. 22 is solved the new perturbation values, $\Delta\psi_j^n$, are added to the old solution vector (Eq. 18) to form the vector of current streamfunction values. After multiplication by the proper weight function the new and old solution vectors are added and the averaged answer is sent to Eq. 17 for use in solving Eq. 23. The same process is then repeated for $\Delta\phi_j^n$. The averaging scheme is then

$$\left.\begin{aligned} \psi_j^n &= W_1(\psi_j^{n-1} + \psi_j^n) + W_2(\psi_j^{n-1}) \\ \phi_j^n &= W_1(\phi_j^{n-1} + \phi_j^n) + W_2(\phi_j^{n-1}) \end{aligned}\right\} \quad \ldots\ldots (25)$$

The repetitive population and solution of Eqs. 22 and 23 using the most current weighted solution vectors, Eq. 25, is the iterative process. Complete specification of the numerical technique closes with a brief description of the method of handling boundary conditions.

Boundary Conditions.—Known nodal boundary values of ψ, $\partial\psi/\partial x$, $\partial\psi/\partial z$, ϕ, $\partial\phi/\partial x$, and $\partial\phi/\partial z$ are introduced into the starting vectors, ψ_j^1 and ϕ_j^1. Since these values never change the corresponding perturbation quantities, $\Delta\psi_j^1$ and $\Delta\phi_j^1$, are always equal to zero. When, as in the case of the variable bottom topography, boundary conditions are specified in terms of normal and tangential derivatives a coordinate transformation of the affected boundary nodes is necessary. The method, readily used in structural analysis, is described in Refs. 4 and 15.

Solution Procedure.—Computation and storage requirements are greatly reduced by dividing the method into two specific programs, an element library program and an iteration program.

The element library creates the stiffness matrices, Φ, θ, λ, π, and ξ (10,12), for each of the "different" or main elements in a particular descretization. Therefore, even if there are 100 elements [Fig. 3(*b*)], stiffness matrices from only four are determined and stored. Although this method encourages element uniformity, the reduction in core storage requirement justifies the approach.

The iterative solution program solves the nonlinear system of equations by the Newton-Raphson procedure. Several organizational features are noteworthy. The solution order of the two equation sets is established by the criterion that the most linear governing equation, Eq. 6 or 7, is solved first. Because the test time per iteration is considerable, no convergence test was incorporated in the computer program. To test for convergence the results from a completed run of n specified iterations were measured against the convergence criterion $|\epsilon| < 0.1\%$. Results not falling between those limits were used as starting vectors and the program run for another specified number of iterations. As compilation time was only 3 sec a considerable economy in computation time resulted. Full advantage is taken of the banded but nonsymmetric slope matrices, S_{ij} and T_{ij}. Once convergence occurs, the solution vectors, ψ_i and ϕ_i, are printed out and passed directly into a card file. These files serve as starting vectors for the next case.

Although an apparently clumsy technique requiring the repeated population of the slope matrices, this method arose in response to the restrictions in computation machinery (all calculations were performed on an IBM 360/65) and the deficiencies in the currently available methods (23,29,32). A brief description of the deficiencies follows.

Other Solution Methods.—The method of Olson (29), the point of departure for this work, is a full Newton-Raphson procedure. With reference to Eqs. 12 and 13, the method sums the element stiffness matrices to global matrices and then applies the Newton-Raphson procedure. The resulting global three-dimensional slope matrices require so much storage that they must be stored on disk and retrieved at each iteration. For this particular problem more time was spent retrieving the global matrices from disk than repeatedly populating S_{ij}^n and T_{ij}^n with the procedure outlined herein.

The method of Skiba (32), an accelerator method, is straightforward. Eqs. 12 and 13 are summed directly to the global system. Using values of ϕ and ψ from the previous iteration the nonlinear terms are evaluated and the algebraic equations are solved for ψ_i^n, not the perturbation quantities. A weighted average of ψ_i^n is formed and used then to solve the system of equations for ϕ_i^n. The process is repeated for the specified number of iterations. The savings in time per iteration is considerable but convergence proceeds at a much slower pace. Again, the storage of large three-dimensional matrices is required.

Finally the method of King (23) eliminates the necessity of using higher order approximations by solving the four conservation equations directly. The iteration technique, a Newton-Raphson procedure, requires solving at once a large matrix for values of the velocities, pressure, and density. Again the very large storage requirements plus the necessity of solving the sensitive pressure equation eliminate this method.

Problem Solutions

The finite element method is now used to solve several example problems.

The results emphasize the convergence, accuracy, and efficiency of the technique. For two of the three problems previous results are used to establish solution accuracy. The third problem, the motivation for the development of the method, is new and thus previous results are not available.

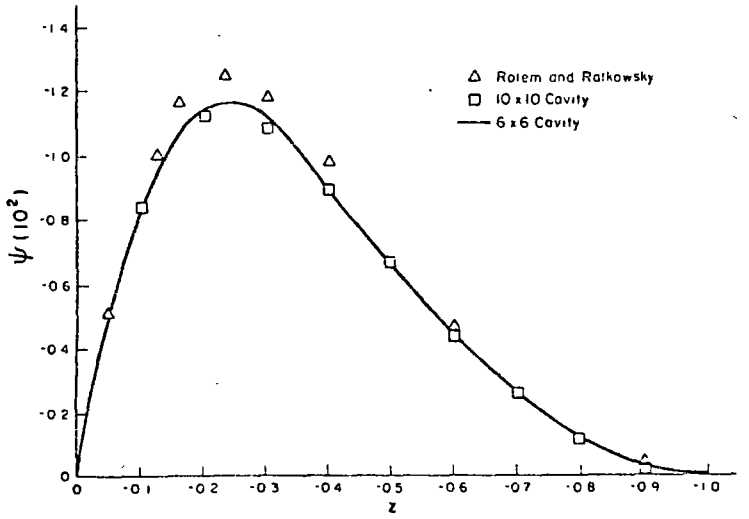


FIG. 4.—Vertical Streamline Profile at $x = 0.3$ and 0.7; Homogeneous, R = 0, Shear Driven, 0(1) Cavity

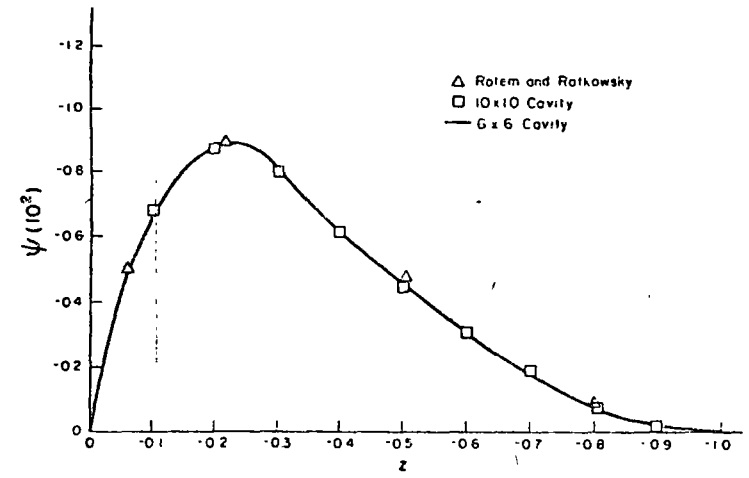


FIG. 5.—Vertical Streamline Profile at $x = 0.5$; Homogeneous, R = 0, Shear Driven, 0(1) Cavity

Linear Biharmonic Model.—The analysis of a shear driven cavity model satisfies three objectives: to determine the number of nodes/elements necessary for accurate cavity results and to determine the effects of a symmetric versus non symmetric flow f discretization.

The problem is defined as follows. A homogeneous water filled cavity is set into motion by the application of a known shear stress. The problem is linearized by considering the R = 0 case. Eq. 6 reduces to $\nabla^4\psi = 0$. The boundary conditions applied to the surface $z = 0$ are $w = 0$, and $\eta_o (\partial u/\partial z) = \tau_x/\rho_o$, in which τ_x is a known shear stress. On the walls and bottom the no-slip condition holds and no flow is allowed through the boundary.

In his work Olson suggests that for a 1 × 1 cavity 72 elements or six on a side is satisfactory. Figs. 4 and 5 then compare the results from a 6 × 6 box [Fig. 3(*a*)] to a 10 × 10 box [Fig. 3(*b*)] and the finite difference results of Rotem and Ratkowsky (30). The streamfunction profiles compare favorably with the published results everywhere but the vortex center. The fullness of the Ratkowsky and Rotem results is not matched by the 6 × 6 cavity and the 10 × 10 cavity due perhaps to the nonconformity adds little refinement. Because mesh refinement failed to improve results substantially the coarser mesh is considered sufficient. By lowering the number of elements from 100 to 72 the system of equations is reduced from 363 to 147, a significant difference.

Natural Convection Model.—The model, utilizing the full nonlinear algorithm, considers fluid motion in a completely enclosed square cavity the sides of which are held at constant but different temperatures. The motion is assumed laminar. A warm temperature, T_h, is evenly applied to the left vertical wall and the right vertical wall is kept at a constant but colder temperature, T_c. If the density is related to the temperature by $\rho = \rho_o [1 - \beta(T - T_o)]$ and if $U = H/\tau$, $T^* = (T - T_c)/(T_h - T_c)$, in which $\bar{\alpha}$ = thermal diffusivity, and τ = reference time = $H^2/\bar{\alpha}$. Then Eqs. 6 and 7 become

$$\left.\begin{aligned} -\mathrm{Pr}\nabla^4\psi + \frac{\partial(\nabla^2\psi,\psi)}{\partial(z,x)} - \mathrm{RaPr}\frac{\partial T}{\partial x} = 0 \\ -\nabla^2 T + \frac{\partial(T,\psi)}{\partial(z,x)} = 0 \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (26)$$

in which Ra = Rayleigh number = $[g\beta(T_h - T_c)/(\bar{\alpha}\eta)]H^3$. The boundary conditions for the flow field are that $u = w = 0$ on the boundary and the temperature field is restricted to an insulated condition on the top and bottom. The negative horizontal temperature gradient, a vorticity source, is governed by the size of the Rayleigh number; therefore, the magnitude of the Prandtl-Rayleigh number product controls the evolution of the solution.

Applied to a 6 × 6 cavity, the procedure begins with a null starting vector for ψ_1^n and seeks a weak heat conduction field for Ra = 1.0. The value of Pr for all runs was 1.0. Computations for Ra up to 10^5 were performed and compare favorably with finite difference solutions by Wirtz (34) and DeVahl Davis (8). The results for Ra = 10^4 and Ra = 10^5 are plotted in Figs. 6 and 7. Streamlines and isotherms for the Ra = 10^5 case are presented in Fig. 8.

The results indicate that the initial choice of the 72-element cavity was accurate. Using this 49-node configuration reduced the problem size considerably, Wirtz used 81 and 289 node configurations and DeVahl Davis used 121 nodes.

The [illegible] differential equation is small. The time [illegible] matrix [illegible] boundaries [illegible] 14[illegible] nations, and form a [illegible]

program solving two systems of algebraic equations per iteration required 106 K of storage and 113.6 sec of IBM 360/65 Fort G execution time. Olson (29) reports that for one iteration on an IBM 360/67, 72 sec of execution time were required to solve one governing equations approximated by 172 algebraic equations. Therefore, the present method represents a fivefold increase in

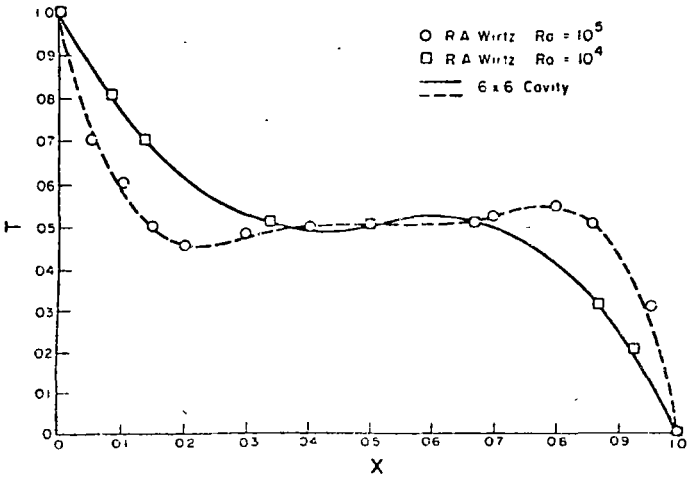


FIG. 6.—Midheight Temperature T Profile; Natural Convection, Ra = 10^4, and 10^5

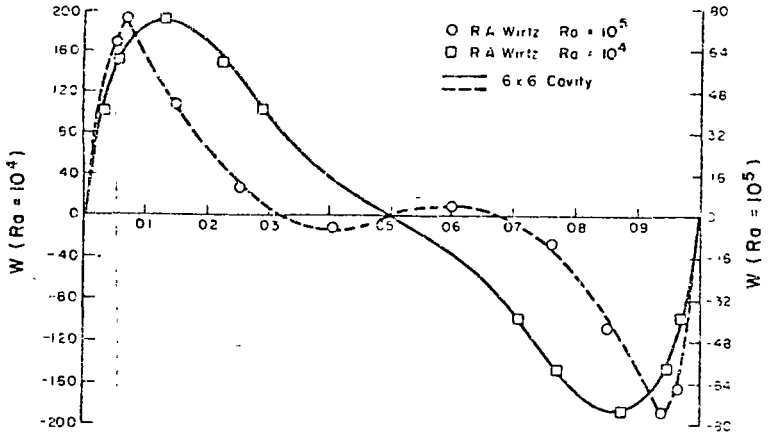


FIG. 7.—Midheight Vertical Velocity w Profile; Natural Convection, Ra = 10^4, and 10^5

efficiency, suggesting that disk storage should be avoided where possible. Chan (6) reports that studies of flow over a cylinder using a 207-equation Newton-Raphson solution took 28.0 sec per equation per iteration. Again the improvement is evident.

convection problem; therefore, the result is not unexpected. Measurement of these properties is done by comparing the numbers of iteration required for convergence and the size of the Raleigh number jump taken while remaining stable. As a step increase of Ra = 10^3 required three to four iterations and no weighted averaging. Step increases of Ra = 10^4 were done in less than nine iterations with simple averaging (i.e., $W_1 = 0.5$, $W_2 = 0.5$). The maximum reliable step size for flow regimes up to Ra = 2×10^5 is 5×10^4 requiring

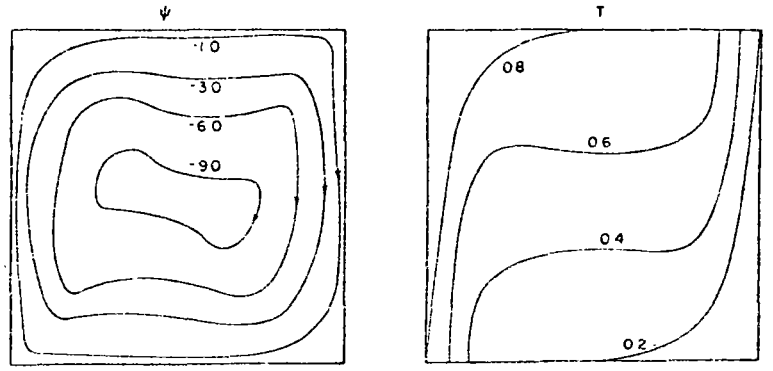


FIG. 8.—Streamline and Isotherm Contours; Natural Convection, Ra = 10^5, Pr = 1.0

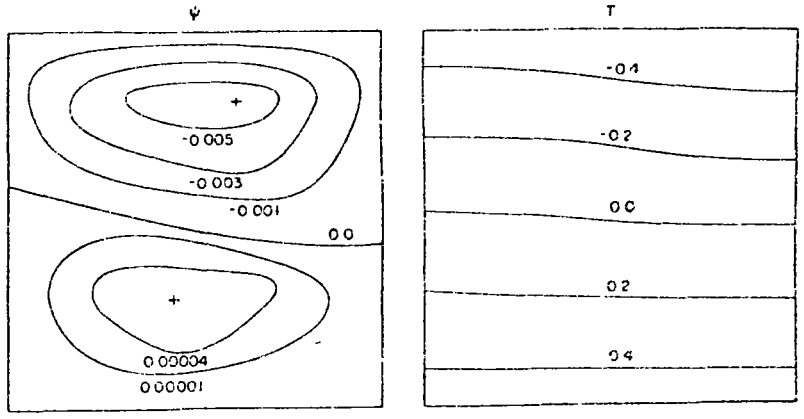


FIG. 9.—Streamline and Isopycnic Contours; Shear Driven, C(1) Cavity, R = 100, Pr = 1.0, Gr = 5,000

10 to 12 iterations with simple averaging. Although successful convergence was achieved at a step size of 7×10^4 the probability of consistently good results is reduced.

De Vahl Davis (8) reports that the stability of the finite difference analog is threatened for step increases greater than 5×10^2. Skiba's (32) method easily handles step increases of Ra = 4×10^4; however, the results are achieved with a Pr = 1,000, a very stable configuration for the problem. No such restriction

is necessary here. Unfortunately, no results are reported on the number of iterations required for Skiba's method.

Combined Convection Problem.—The primary motivation for the algorithm development was the solution of a combined convection problem. Since results are not available for this problem rigorous checking of the method against the established results of the previous cases indicates the correctness of the solution. The problem is this section explores the interplay of wind shear and a stable vertical density gradient, and is peculiar in that distinct circulation "cells" can occur.

The cavity model considers fluid motion in a cavity founded on the sides and bottom and set into motion by a known applied shear stress on the surface. A stable vertical temperature gradient is imposed by specifying maximum and minimum temperatures on the top and bottom of the cavity. The surface temperature is T_T and the bottom temperature is T_B and if U (Eq. 5) is $\tau_x H/(\eta_o \rho_o)$ then Eqs. 6 and 7 become

$$\left.\begin{aligned} -\frac{\eta}{R}\nabla^4\psi + \frac{\partial(\nabla^2\psi,\psi)}{\partial(z,x)} - \frac{Gr}{R^2}\frac{\partial T}{\partial x} = 0 \\ -\frac{E}{PrR}\nabla^2 T + \frac{\partial(T,\psi)}{\partial(z,x)} = 0 \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (27)$$

The boundary conditions are that $u = w = 0$ on the bottom and sides and $w = 0$ on the surface. Further $\partial u/\partial z = \tau_x/(\eta_o \rho_o)$ is specified on the surface. Heat flux is prohibited through the sides.

Calculations were performed for a variety of flow conditions. The range of parameters was: R of 1 to 1,000, Pr of 1 to 10, and Gr of 1 to 10^5. Extensive results to this problem will appear in another paper, however, a streamfunction and isopycnal plot is presented here, in Fig. 9 for demonstration purposes.

Iteration and convergence behavior is the same as the natural convection problem.

Conclusions

From this study several conclusions about the use and application of this method are drawn.

The full Newton-Raphson method can be made an efficient and, for this class of problem, a preferred computational scheme. With the proper use of element libraries the necessity of using disk storage is eliminated. The use of higher order interpolations reduces the number of necessary elements and is highly recommended.

This paper establishes that the use of the FEM encourages solution stability and iterative speed. The method is preferred over finite difference schemes for this class of steady problems. Problems and questions are plentiful and unfortunately the solution of coupled nonlinear equations by the FEM remains unresearched even in light of recent Swansea (22) conference. In physical domains where analytical answers are impossible the FEM seems to be the correct approach.

Acknowledgment

Work described here was supported in part by the National Science Foundation under Research Grant GK-23992. Work done by the senior writer at Cornell University was supported by an Environmental Protection Agency Research Fellowship. The invaluable assistance of R. Gallagher is gratefully acknowledged.

Appendix I.—References

1. Argyris, J. H., Mareczek, C., and Scharpf, D., "Two and Three-Dimensional Flow Using Finite Elements," *Aeronautical Journal*, Vol. 73, Nov., 1969, pp. 961-964.
2. Atkinson, B., et al., "Low Reynolds Number Developing Flow," *American Institute of Chemical Engineers*, Vol. 15, July, 1969.
3. Bazely, G. P., et al., Triangular Elements in Bending-Conforming and Nonconforming Solutions," *Proceedings of the Conference on Matrix Methods in Structural Mechanics*. Air Force Institute of Technology, Wright Patterson Air Force Base, Ohio, 1965.
4. Bedford, K. W., "A Numerical Investigation of Stably Stratified, Wind Driven Cavity Flow by the Finite Element Method," thesis presented to the Cornell University, at Ithaca, N Y., in 1974, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.
5. Bird, R. B., Steward, W. E., and Lightfoot, E. N., *Transport Phenomena*, John Wiley and Sons, Inc., New York, N.Y., 1960.
6. Chan, S. T. K., and Larock, B. E., "Potential Flow Around a Cylinder Between Parallel Walls by Finite Element Methods," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 98, No. EM5, Proc. Paper 9228, Oct., 1972, pp. 1317-1322.
7. Cheng, R., "Numerical Investigation of Lake Circulation Around Islands by the Finite Element Method," *International Journal for Numerical Methods in Engineering*, Vol 5, No. 1, 1972.
8. De Vahl Davis, G., "Laminar Natural Convection in an Enclosed Retangular Cavity," *International Journal of Heat and Mass Transfer*, Vol. 11, 1968, p. 1675.
9. De Vries, G., and Norrie, D. H., "The Application of the Finite Element Technique to Potential Flow Problems," *Journal of Applied Mechanics*, Paper No. 71-APM-22, 1971.
10. Doctors, L. J., "An Application of the Finite Element Technique to Boundary Value Problems of Potential Flow," *International Journal of Numerical Methods in Engineering*, Vol. 2, 1970, p. 243.
11. Donnelly, R. J., Horman, R., and Prigogine, I., eds., *Non Equilbrium Thermodynamics, Variational Techniques, and Stability*, University of Chicago Press, Chicago, Ill., 1966.
12. Felippa, C. A., "Refined Finite Element Analysis of Linear and Nonlinear Two Dimensional Structures," *Report SESM 66-22*, Structural Engineering Laboratory, University of California, Berkeley, Calif., Oct., 1966.
13. Finlayson, B. A., *The Method of Weighted Residuals and Variational Principles*, Vol. 87, Academic Press, New York, N.Y., 1972.
14. Finlayson, B. A., and Scriven, L. E., "On the Search for Variational Principles," *Heat and Mass Transfer*, Vol. 10, 1967, p. 799-821.
15. Gallagher, R., and Chan, S. T. K., "Higher Order Finite Element Analysis of Lake Circulation," *Proceedings of the Conference on Computers in Fluid Dynamics*, Analysis and Design, Polytechnic Institute of Brooklyn, Brooklyn, N.Y., 1973.
16. Gallagher, R., Liggett, J. A., and Chan, S. T. K., "Finite Element Circulation Analysis of Variable-Depth Shallow Lakes," *Journal of the Hydraulics Division*, ASCE, Vol. 99, No. HY7, Proc. Paper 9855, July, 1973, pp. 1083-1096.
17. Gebhart, B., *Heat Transfer*, McGraw Hill Book Co., Inc., New York, N.Y., 1971.
18. Glansdorf, P., and Prigogine, I., "Sur les proprietes differntielles de la production d'entropie," *Physica Grav.*, Vol. 20, 1954, p. 773.
19. sdorf, P., and Prigogine, I., "On a General Evolution Criterion in Macroscopics," *Physica Grav.*, Vol. 30, 1964, p.351.

20 Gurymoo... G. L., "Application of the Finite Element Method for Simulation of Surface Water Transport Problems," Institute of Water Resources, University of Alaska, *Report No. IWR 21*, College, Alaska, June, 1972.
21. Ikenouchi, M., and Kimura, N., "An Approximate Numerican Solution of the Navier Stokes Equations by the Galerkin Method," *Finite Elements in Flow Problems*, J. T. Oden, ed., *Conference Proceedings*, Swansea, Wales, 1974.
22. Kawahara, M., Yoshimura, N., and Nakagawa, K., "Analysis of Steady Incompressible Viscous Flow," *Finite Elements in Flow Problems*, J. T. Oden, ed., *Conference Proceedings*, Swansea; Wales, 1974.
23. King, I. P., Norton, W. R., and Orlob, G. T., "A Finite Element Solution for Two-Dimensional Stratified Flow," Water Resources Engineers, Inc., Walnut Creek, Calif., Apr., 1973.
24. Lardner, T. J., "Biots Variational Principle in Heat Conduction," *American Institute of Acronautics and Astronautics Journal*, Vol. 1, 1963, p. 196.
25. Lomax, H., "Numerical Solution of Partical Differential Equations Governing Convection," *Technical Report*, North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development, 1970.
26. Oden, J. T., *Finite Elements of Nonlinear Continua*, McGraw-Hill Book Co., Inc., New York, N.Y., 1972.
27. Oden, J. T., and Sornogyi, D., "Finite Element Applications in fluid Dynamics," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 95, No. EM3, Proc. Paper 6584, June, 1969, pp. 821-826.
28. Oden, J. T., et al., eds., *Finite Element Methods in Flow Problems, Proceedings, of the International Symposium on Finite Element Methods in Flow Problems*, Swansea, Wales, Jan., 1972.
29. Olson, M. D., "A Variational Finite Element Method for Two Dimensional Steady Viscous Flows," *Proceedings of McGill-EIC* Conference on Finite Element Methods in Civil Engineering, Montreal, Canada, June 1-2, 1972.
30. Ratkowsky, D. A., and Rotem, Z., "Viscous Flow in a Rectangular Cut," *Physics of Fluids*, Vol. 11, 1968, p. 2761.
31. Schlicting, H., *Boundary Layer Theory*, McGraw-Hill Book Co., Inc., New York, N.Y., 1968.
32. Skiba, E., Unny, T. E., and Weaver, D. S., "A Finite Element Solution for a Class of Two Dimensional Viscous Fluid Dynamics Problems," *Computer Aided Engineering, Proceedings of a Symposium*, University of Waterloo, Waterloo, Ontario, Canada, 1971.
33. Tong, P., "The Finite Element Method for Fluid Flow," Paper U.S. 5-4, Japan-United States Seminar on Matrix Methods of Structural Analysis and Design, Tokyo, Japan, Aug., 1969.
34. Wirtz, R. A., Briggs, D. G., and Chen, C. F., "Physical and Numerical Experiments on Layered Convection in a Density Stratified Fluid," *Geophysical Fluid Dynamics*, Vol. 3, 1962, pp. 265-288.
35. Zienkiewicz, O. C., *The Finite Element Method in Engineering Science*, McGraw-Hill Book Co. Ltd., London, England, 1971.

Appendix II.—Notation

The following symbols are used in this paper:

A, A_i = element area and area of subregion i, respectively;
a_i, b_i, c_i = element geometry coefficients;
Co, Cx, Cz
Cxx, Cxz, Czz = coefficient matrices for shape functions and first and second derivatives;
D_1, D_2 = operator notation for governing equations;
d, e, f = column vectors of element integrable functions;
E, Eo = turbulent density eddy viscosity and reference values;

F_i^n, f_i^n = ith global and element algebraic equations at nth iteration for vorticity equations;
G_i^n, g_i^n = ith global and element algebraic equation at nth iteration for conservation equation;
Gr = Grashof number;
g = gravitational coefficient;
H = maximum cavity depth;
i, j, k = indices;
L = cavity length;
L_1, L_2, L_3 = element area coordinates;
N = maximum number of iterations;
N_i = element shape functions;
P_i = boundary condition vector for vorticity element stiffness equations;
Pr = Prandtl number;
p = pressure;
Q_i^e = boundary condition vector for conservation element stiffness equations;
R = Reynolds number;
Ra = Raleigh number;
R_i^* = residual for ith equation;
S_{ij}^n = ith global slope matrix at nth iteration for vorticity equation;
T = temperature;
To, Th, Tc = reference, left wall and right wall temperatures;
T_{ij}^n = ith global slope matrix at nth iteration for conservation equation;
U, W = reference velocities;
u, w = horizontal and vertical velocities;
W_1, W_2 = weight function;
X, Z = rectangular Cartesian coordinates;
α = diffusivity for conservation quantity ϕ;
$\bar{\alpha}$ = molecular thermal difusivity at reference density ρ_o;
β = volumetric expansion coefficient;
ϵ = error limit;
η, η_o = eddy viscosity and reference value;
$\theta_{ikj}, \lambda_{ij}$,
$\Phi_{ij}, \xi_{ij}, \pi_{ikj}$ = element stiffness matrices;
ρ, ρ_o = density and reference value;
ρ_T, ρ_B = density at the cavity top and bottom;
τ = thermal diffusion time;
τ_x = surface shear stress;
$\phi, \bar{\phi}$ = conservation quantity and series representation;
$\phi_i^n, \Delta\phi_i^n$ = vectors of ϕ and perturbation quantities at nth iteration;
$\psi, \bar{\psi}$ = streamfunction and its series representation; and
$\psi_i^n, \Delta\psi_i^n$ = vectors of ψ and perturbation quantities at nth iteration.

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

USOS DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPRESION EN AGUAS, LAGOS Y RIOS

THE CALCULUS OF VARIATIONS

MAYO,1978.

Palacio de Minería Calle de Tacuba 5, primer piso México 1, D. F. Tel.: 521-40-20 5 Líneas

THE CALCULUS OF VARIATIONS

1.1 Introduction

In the study of elementary differential calculus we investigate how certain quantities say 'f', varies as another quantity, say 'x' is altered when there is a relationship between x and ϕ. We say that f is a function of x, $f = f(x)$, if there is some rule whereby we can calculate the value of ϕ if we know the value of x. One particular investigation we make is into the determination of the maximum and minimum values 'ϕ' may have and the conditions for finding these values.

In the calculus of variations we consider a similar type of problem, we study how a quantity called a 'functional' varies as we change the function $f = f(x)$ to another function, say $\varphi = \varphi(x)$, and in particular try to find the function which gives the functional an extremum (maximum or minimum) value.

A functional is a quantity whose value depends upon a function, for example

$$F = \int_0^1 f(x)\, dx \qquad (1.1)$$

is a simple functional. The value of the definite integral will depend upon which function $f(x)$ we chose. We shall be concerned with integrals in the form of definite integrals in which the integrand may not only include the function f but also its derivatives. A typical problem might be to find the function $f(x)$ which gives

$$F = \int_0^L \left\{ \left(\frac{d^2 f}{dx^2}\right)^2 - 2a f \right\} dx \qquad (1.2)$$

a minimum value, and also satisfies some specified conditions at $x = 0$ and $x = L$. Further restrictions will also be placed upon the range of functions out of which $f(x)$ is to be selected, these generally will require that f is a continuous function of x and that some of its derivatives are also continuous. The range of functions which satisfy the boundary conditions and have the required degree of continuity we call 'admissible' functions.

We will denote a functional by $F(f)$ if f is the required function. If more than one function is required we list these in the parenthesis, e.g. $F(u,v,w)$ means that we will be looking for functions u, v and w. The integrand we write as $I(\ )$, and in the parenthesis we write much of the required functions and their derivatives as appear in the integrand and the independent variables. Thus the functional cited above -(1.2)- would be written as

$$F(f) = \int_0^L I(f, \frac{d^2f}{dx^2}, x)\, dx$$

or

$$F(f) = \int_0^L I(f, f_{xx}, x)\, dx \qquad (1.3)$$

where the suffix notation indicates differentiation

$$f_x = \frac{df}{dx} \qquad f_{xx} = \frac{d^2f}{dx^2} \quad \text{, etc.}$$

The first problem we shall consider is the determination of the function that gives some integral which depends upon f and f_x a maximum, minimum or more generally a stationary value.

1.2 Functionals of the form $F(f) = \int_{x_1}^{x_2} I(f, f_x, x)dx$.

Let us try to find the function $f(x)$ that has the values $f_1 = f(x_1)$ at $x = x_1$ and $f_2 = f(x_2)$ at $x = x_2$ and gives a stationary value to the functional.

$$F(f) = \int_{x_1}^{x_2} I(f, \frac{df}{dx}, x)dx \qquad (1.5)$$

We will use the symbol f to denote the function that gives the stationary value to this functional. Any other function which passes through the points (f_1, x_1) and (f_2, x_2) can be put in the form $f(x) + \eta(x)$, where $\eta(x)$ is zero at $x = x_1$ and $x = x_2$ (figure (1.1).

The expression,

$$h(x, \epsilon) = f(x) + \epsilon\, \eta(x) \qquad (1.6)$$

will represent a series of curves each of which passes through the specified end points. We limit ourselves to functions $f(x)$ and $\eta(x)$ which are continuous in the interval $x_1 < x < x_2$. For each function $\eta(x)$ the above expression $h(x, \epsilon)$ will give a family of curves, and all the possible such families will contain $f(x)$ when ϵ is zero. If we now evaluate the functional $F(f)$ we know that this will be an extremum when $\epsilon = 0$ and its value will then be $F(y)$. The value of $F(f)$ will vary with ϵ and we know, from the definition that f extremizes the functional, that $F(f)$ will be an extremum with respect to ϵ when $\epsilon = 0$.

Now

$$F(h) = \int_{x_1}^{x_2} I(h, h_x, x)dx \qquad (1.7)$$

and

$$\frac{dh}{dx} = \frac{df}{dx} + \epsilon \frac{d\eta}{dx} = f_x + \epsilon\, \eta_x \qquad (1.8)$$

If we differentiate (1.7) with respect to ϵ,

$$\frac{dF(h)}{d\epsilon} = \int_{x_1}^{x_2} \frac{dI}{d\epsilon}\, dx = \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial h}\frac{\partial h}{\partial \epsilon} + \frac{\partial I}{\partial h_x}\frac{\partial h_x}{\partial \epsilon}\right) dx \quad (1.9)$$

(Note that x does not vary).

The second term arises as not only does I change with variations in h(x) but also will variations in dh/dx (or h_x).

From $h = f + \epsilon\eta$ and $h_x = f_x + \epsilon\eta_x$ (1.10)

we find

$$\frac{\partial h}{\partial \epsilon} = \eta \quad \text{and} \quad \frac{\partial h_x}{\partial \epsilon} = \eta_x \quad (1.11)$$

Whence

$$\frac{dF(h)}{d\epsilon} = \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial h}\eta + \frac{\partial I}{\partial h_x}\eta_x\right) dx \quad (1.12)$$

Integrating the second term in the right hand side by parts we get

$$\frac{dF(h)}{d\epsilon} = \int_{x_1}^{x_2} \left\{\frac{\partial I}{\partial h}\eta - \frac{d}{dx}\left(\frac{\partial I}{\partial h_x}\right)\eta\right\} dx + \left|\frac{\partial I}{\partial h_x}\eta\right|_{x_1}^{x_2} \quad (1.13)$$

The last term in this expression is zero as $\eta(x)$ is zero at both limits $x = x_1$ and $x = x_2$. We thus find,

$$\frac{dF(h)}{d\epsilon} = \int_{x_1}^{x_2} \left\{\frac{\partial I}{\partial h} - \frac{d}{dx}\left(\frac{\partial I}{\partial h_x}\right)\right\}\eta\, dx \quad (1.14)$$

For a stationary value of F(f) this equation must be zero, and this occurs when $\epsilon = 0$ and $h_x = f_x$. We thus have,

$$\left.\frac{dF(h)}{d\epsilon}\right|_{\epsilon=0} = \int_{x_1}^{x_2} \left\{\frac{\partial I}{\partial f} - \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right)\right\}\eta\, dx = 0 \quad (1.15)$$

This is to be true for all the admissible functions η, that is for all functions h(x) which are zero for $x_1 = x$ and $x = x_2$ and are contin-

uous in the interval $x_1 < x < x_2$.

The Basic Lemma of the Calculus of Variations

"If

$$\int_{x_1}^{x_2} \alpha(x)\, \eta(x)\, dx = 0 \qquad (1.16)$$

for all admissible functions $\eta(x)$ then $\alpha(x) = 0$ throughout the interval."

Let $\alpha(x)$ not be zero at some point $x = \bar{x}$. Since all the necessary functions are assumed to be continuous, $\alpha(x)$ will have the same sign in small interval $a = \bar{x} - \xi < \bar{x} < \bar{x} + \xi = b$. For instance (figure 1.2) the function

$$\begin{aligned} \eta(x) &= 0 \quad , \quad x < a \quad , \quad x > b \\ \eta(x) &= (x-a)^2(x-b)^2 \quad , \quad a < x < b \end{aligned} \qquad (1.17)$$

will be continuous and have a continuous derivative, and is thus admissible. Since in the interval $a < x < b$, $\eta(x)$ is essentially positive the integral,

$$\int_{x_1}^{x_2} \alpha(x)\, \eta(x) dx = \int_{a}^{b} \alpha(x)\, \eta(x) dx \qquad (1.18)$$

will not be zero, which contradicts our initial condition. Hence $\alpha(x)$ cannot have any value at any point in the region considered.

A similar argument applies if we require the nth derivative of $\eta(x)$ to be continuous, but in this case we use

$$\begin{aligned} \eta(x) &= 0 \quad , \quad x < a \, , \quad x > b \\ \eta(x) &= (x-a)^{n+1}(x-b)^{n+1} \quad , \quad a < x < b \end{aligned} \qquad (1.19)$$

since this will give the required continuity.

Thus, the function $f(x)$ which gives a stationary value for F - equation (1.15) - is such that satisfies $\alpha(x) \equiv 0$ or,

$$\frac{\partial I}{\partial f} - \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right) = 0 \qquad (1.20)$$

throughout the interval $x_1 < x < x_2$

Equation (1.20) is called the Euler-Lagrange equation and represents a second order differential equation.

Example 1.1 Find the condition that

$$F(f) = \int_0^1 \left\{A\left(\frac{df}{dx}\right)^2 + B\left(\frac{df}{dx}\right) + Cf^2 + Df\right\} dx \qquad (a)$$

should be stationary, subjected to $f(1) = 0$, $f(o) = 0$

We have,

$$I = A\left(\frac{df}{dx}\right)^2 + B\left(\frac{df}{dx}\right) + Cf^2 + Df \qquad (b)$$

Thus,

$$\frac{\partial I}{\partial f} = 2Cf + D \qquad \frac{\partial I}{\partial f_x} = 2Af_x + B \qquad (c)$$

Whence, equation (1.20) becomes,

$$\frac{\partial I}{\partial f} - \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right) = 2Cf + D - 2Af_{xx} = 0 \qquad (d)$$

or

$$2A\frac{d^2f}{dx^2} - 2Cf - D = 0 \qquad (e)$$

with boundary conditions $f = 0$ at $x = o$, $x = 1$.

Example 1.2 The deflection of a strip of film, with film tension T, under normal pressure p is such that it extremizes the following functional

$$F(w) = \int_0^L \left\{ \frac{T}{2}\left(\frac{dw}{dx}\right)^2 - pw \right\} dx \qquad (a)$$

where w is the deflection and is zero at the ends of the interval, $x = o, L$.

Now,

$$I = \frac{T}{2}\left(\frac{dw}{dx}\right)^2 - pw \qquad (b)$$

$$\frac{\delta I}{\delta w} = -p \qquad \frac{\delta I}{\delta w_x} = Tw_x \qquad (c)$$

Whence the Euler-Lagrange condition gives

$$-p - \frac{d}{dx}(Tw_x) = 0 \qquad (d)$$

or

$$\frac{d^2w}{dx^2} + \frac{p}{T} = 0 \qquad (e)$$

The solution of this equation is,

$$w = -\frac{p}{2T}x^2 + Ax + B \qquad (f)$$

and since $w(o) = 0$, $w(L) = 0$

$$w = \frac{p}{2T}x\,(L - x) \qquad (g)$$

1.3 Functionals of the Form $F(f) = \int_{x_1}^{x_2} I(f, f_x, f_{xx}, x)\,dx$

We consider now functionals whose value depends upon the nature of a function and its first derivatives

$$F(f) = \int_{x_1}^{x_2} I(f, f_x, f_{xx}, x)\,dx \qquad (1.21)$$

The required value is the function $f(x)$ that gives a stationary value to this function and has specified values f_1, f_2 at $x = x_1$, x_2, and also specified values for the first derivative. $f_x(x_1) = f_{x_1}$ and $f_x(x_2) = f_{x_2}$

We proceed as before, denoting the energizing function by $\phi(x)$, and considering functions $h(x) = f(x) + \epsilon\eta(x)$. In this it is necessary that $\eta(x)$ is zero at $x = x_1$ and $x = x_2$ and that the first derivative

$\eta(x)$ is zero at $x = x_1$ and $x = x_2$ and that the first derivative $\eta_x(x)$ is also zero at the ends of the interval (figure 1.3). Noting that,

$$h = h + \epsilon\eta \quad , \quad \frac{\partial h}{\partial \epsilon} = \eta$$

$$h_x = f_x + \epsilon\eta_x \quad , \quad \frac{\partial h_x}{\partial \epsilon} = \eta_x \tag{1.22}$$

$$h_{xx} = f_{xx} + \epsilon\eta_{xx} \quad , \quad \frac{\partial h_{xx}}{\partial \epsilon} = \eta_{xx}$$

we find that the value of a functional $F(h)$ is

$$F(h) = \int_{x_1}^{x_2} I(h, h_x, h_{xx}, x)dx \tag{1.23}$$

and this is stationary with respect to ϵ when $\epsilon = 0$. We have

$$\frac{\partial F}{\partial \epsilon} = \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial h}\frac{\partial h}{\partial \epsilon} + \frac{\partial I}{\partial h_x}\frac{\partial h_x}{\partial \epsilon} + \frac{\partial I}{\partial h_{xx}}\frac{\partial h_{xx}}{\partial \epsilon}\right)dx =$$

$$= \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial h}\eta + \frac{\partial I}{\partial h_x}\eta_x + \frac{\partial I}{\partial h_{xx}}\eta_{xx}\right)dx \tag{1.23}$$

and since when $\epsilon = 0$, $f(x) = \phi(x)$,

$$\left.\frac{\partial F}{\partial \epsilon}\right|_{\epsilon=0} = \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial f}\eta + \frac{\partial}{\partial f_x}\eta_x + \frac{\partial}{\partial f_{xx}}\eta_{xx}\right) dx = 0 \tag{1.24}$$

We now integrate the second term by parts once, and the third term by parts twice, i.e.

$$\int_{x_1}^{x_2} \frac{\partial I}{\partial f_{xx}}\eta_{xx}\,dx = \left|\frac{\partial I}{\partial f_{xx}}\eta_x\right|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right)\eta_x\,dx =$$

$$= \left|\frac{\partial I}{\partial f_{xx}}\eta_x\right|_{x_1}^{x_2} - \left|\frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right)\eta\right|_{x_1}^{x_2} + \int_{x_1}^{x_2} \frac{d^2}{dx^2}\left(\frac{\partial I}{\partial f}\right)\eta\,dx \tag{1.25}$$

We can write (1.24) as

$$\int_{x_1}^{x_2} \{ \frac{\delta I}{\delta f} - \frac{d}{dx}(\frac{\delta I}{\delta f_x}) + \frac{d^2}{dx^2}(\frac{\delta}{\delta f_{xx}}) \} \eta \, dx + \{ \frac{\delta I}{\delta f_x} - \frac{d}{dx}(\frac{\delta}{\delta f_{xx}}) \} \eta \Big|_{x_1}^{x_2} +$$

$$+ \frac{\delta I}{\delta f_{xx}} \eta_x \Big|_{x_1}^{x_2} = 0 \qquad (1.26)$$

But the admissible functions $\eta(x)$ are such that $\eta(x) = 0$ and $\eta_x(x) = 0$ at $x = x_1$ and $x = x_2$, so that all the limit terms in (1.26) are zero. The integral term is zero for all $\eta(x)$, hence by the Basic Lemma,

$$\frac{\delta I}{\delta f} - \frac{d}{dx}(\frac{\delta I}{\delta f_x}) + \frac{d^2}{dx^2}(\frac{\delta}{\delta f_{xx}}) = 0 \qquad (1.27)$$

This is the Euler equation corresponding to functional (1.21) and represents a fourth order differential equation for the function f.

Example 1.3 Find the condition that,

$$\int_0^L \{ \frac{EI}{2} (\frac{d^2w}{dx^2})^2 + \frac{k}{2} w^2 - pw \} \, dx \qquad (a)$$

shall be a minimum. The function w and its derivative having specified values at $x = o$ and $x = L$.

$$I = \frac{EI}{2} (\frac{d^2w}{dx^2})^2 + \frac{k}{2} w^2 - pw \qquad (b)$$

Thus

$$\frac{\delta I}{\delta w} = kw - p \quad , \quad \frac{\delta I}{\delta w_x} = 0 \quad , \quad \frac{\delta I}{\delta w_{xx}} = EI \frac{d^2w}{dx^2} \qquad (c)$$

whence,

$$\frac{d^2}{dx^2} (EI \frac{d^2w}{dx^2}) + kw - p = 0 \qquad (d)$$

which is the equilibrium equation for a beam on elastic foundation.

1.4 Functionals involving derivatives up to the n-th degree.

Let us consider now the problem of finding the condition that the function $f(x)$ gives a stationary value to the functional

$$F(f) = \int_{x_1}^{x_2} I(f, f_x, f_{xx}, \ldots, f_{x(n)}, x)dx \qquad (1.28)$$

where $f_{x(n)}$ denotes $\frac{d^n f}{dx^n}$. The function f and its first $(n-1)$ derivatives have some specified values at the limits $x = x_1$ and $x = x_2$. The admissible functions can again be written in the form,

$$h(x) = f(x) + \epsilon\,\eta(x) \qquad (1.29)$$

$f(x)$ being the required function giving the stationary value to $F(f)$, but now $\eta(x)$ and its first$(n-1)$ derivatives must be zero at the ends of the interval of integration. We proceed as before and find,

$$\frac{dF(h)}{d\epsilon} = \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial h}\frac{\partial h}{\partial \epsilon} + \frac{\partial I}{\partial h_x}\frac{\partial h_x}{\partial \epsilon} + \frac{\partial I}{\partial h_{xx}}\frac{\partial h_{xx}}{\partial \epsilon} + \ldots + \frac{\partial I}{\partial h_{x(n)}}\frac{\partial h_{x(n)}}{\partial \epsilon}\right)dx = 0 \qquad (1.30)$$

and with $\epsilon = 0$ (that is $h = f$, $h_x = f_x$, etc),

$$\frac{dF(h)}{\partial \epsilon} = \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial f}\eta + \frac{\partial I}{\partial f_x}\eta_x + \frac{\partial I}{\partial f_{xx}}\eta_{xx} + \ldots + \frac{\partial I}{\partial f_{x(n)}}\eta_{x(n)}\right)dx = 0 \qquad (1.31)$$

We now integrate the various terms by parts until in each case the second term in the integrand becomes η. Thus a typical term

$$\int_{x_1}^{x_2} \frac{\partial I}{\partial f_{x(j)}}\eta_{x(j)}dx = \left|\frac{\partial I}{\partial f_{x(j)}}\eta_{x(j-1)}\right|_{x_1}^{x_2} - \left|\frac{d}{dx}\frac{\partial I}{\partial f_{x(j)}}\eta_{x(j-2)}\right|_{x_1}^{x_2} +$$

$$+ \left|\frac{d^2}{dx^2}\frac{\partial I}{\partial f_{x(j)}}\eta_{x(j-3)}\right|_{x_1}^{x_2} \ldots + (-1)^{j-1}\left|\frac{d}{dx^{j-1}}\frac{\partial I}{\partial f_{x(j)}}\eta_x\right|_{x_1}^{x_2} + \qquad (1.32)$$

$$+ (-1)^j \int_{x_1}^{x_2} \frac{d^j}{dx^j}\frac{\partial I}{\partial f_{x(j)}}\eta\,dx$$

All the terms to be evaluated at the limits are zero since they all include factors of η or its first n-1 derivatives. The remaining integrals each contain the factor $\eta(x)$, and since $\eta(x)$ is arbitrary, apart from its continuity conditions, the condition for stationary value becomes,

$$\frac{\partial I}{\partial f} - \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right) + \frac{d^2}{dx^2}\left(\frac{\partial I}{\partial f_{xx}}\right) \cdots + (-1)^j \frac{d^j}{dx^j}\left(\frac{\partial}{\partial f_{x(j)}}\right) \cdots$$

$$+ (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial I}{\partial f_{x(n)}}\right) = 0 \qquad (1.33)$$

This condition represents, in most cases, a differential equation of order 2n.

Example 1.4 Find the condition that

$$F(f) = \int_{x_1}^{x_2} \left\{ \frac{A}{2}\left(\frac{d^4 f}{dx^4}\right)^2 + \frac{B}{2}\left(\frac{d^2 f}{dx^2}\right)^2 + C\left(\frac{df}{dx}\right)^2 + Df^2 + Ef \right\} dx \qquad (a)$$

shall be a minimum, f and its first three derivatives having specified values at $x = x_1, x_2$.

$$\frac{\partial I}{\partial f} = 2Df + E \quad , \quad \frac{\partial I}{\partial f_x} = 2C\frac{df}{dx} \quad , \quad \frac{\partial I}{\partial f_{xx}} = B\frac{d^2 f}{dx^2} \qquad (b)$$

$$\frac{\partial I}{\partial f_{x(3)}} = 0 \quad , \quad \frac{\partial I}{\partial f_{(4)}} = A\frac{d^4 f}{dx^4}$$

Whence the Euler equation becomes,

$$2Df + E - 2C\frac{d^2 f}{dx^2} + B\frac{d^4 f}{dx^4} + A\frac{d^8 f}{dx^8} = 0 \qquad (c)$$

1.5 Boundary Conditions

In the previous section we have considered that the required function, and its first n-1 derivatives were specified at each end of the region of integration. Let us look now at the problem: 'find the function f(x) which gives a stationary value to'

$$F(f) = \int_{x_1}^{x_2} I(f, f_x, x)dx \qquad (1.34)$$

with the condition $f(x_1) = f_1$ and the value of $f(x)$ at x_2 being unspecified. If we proceed as before the admissible function $\eta(x)$ must be such that $\eta(x_1) = 0$, but there is no restriction placed upon $\eta(x_2)$. As before we find (equation (1.13) with $\epsilon = 0$, whence $f = \phi$).

$$\int_{x_1}^{x_2} \left\{ \frac{\partial I}{\partial f}\eta - \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right)\eta \right\}dx + \left| \frac{\partial I}{\partial f_x}\eta \right|_{x_2} = 0 \qquad (1.35)$$

The admissible functions include those for which $\eta(x_2) = 0$, so that we can say that the integral term must be zero for all admissible functions with $\eta(x) = 0$ at $x = x_1$, and $x = x_2$, so that the condition

$$\frac{\partial I}{\partial f} - \frac{d}{dx}\frac{\partial I}{\partial f_x} = 0 \qquad (1.36)$$

still holds. But the range of admissible functions also includes functions $\eta(x)$ which are not zero at $x = x_2$, whence if

$$\left| \frac{\partial I}{\partial f_x}\eta \right|_{x_2} = 0 \qquad (1.37)$$

we must have

$$\frac{\partial I}{\partial f_x} = 0 \qquad \text{at } x = x_2$$

Thus of all the continuous functions that pass through $f(x_1) = f_1$, the one that gives a stationary value to $F(f)$ will satisfy the differential equation (1.36) and also the boundary condition (1.38) at $x = x_2$. The boundary condition (1.38) is called the 'natural boundary condition' of the problem.

Note that by adding terms to the variational functional it is possible to alter the 'natural' boundary conditions but the Euler equation will remain the same. Let us add the functions H_1 valid at x_1 and H_2 valid at x_2 to the functional (1.7).

$$F(h) = \int_{x_1}^{x_2} I(h, h_x, x)dx + H_2(f) + H_1(f) \qquad (1.39)$$

If we differentiate (1.39) with respect to ϵ and substitute h by (1.6), we obtain

$$\left.\frac{dF}{d\epsilon}\right|_{\epsilon=0} = \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial f} - \frac{d}{dx}\frac{\partial I}{\partial f_x}\right)\eta\, dx + \left.\frac{\partial I}{\partial f_x} + \frac{\partial H_2}{\partial f}\right|_{x_2}\eta - \left.\frac{\partial I}{\partial f_x} - \frac{\partial H_1}{\partial f}\right|_{x_1}\eta \qquad (1.40)$$

Thus the Euler equation is the same as obtained before but the natural boundary conditions are,

$$\frac{\partial I}{\partial f_x} - \frac{\partial H_1}{\partial f} = 0 \text{ at } x_1 \quad , \quad \frac{\partial I}{\partial f_x} - \frac{\partial H_2}{\partial f} = 0 \text{ at } x_2 \qquad (1.41)$$

Example 1.5 Find the function which extremises

$$F(f) = \int_0^1 \left\{\left(\frac{df}{dx}\right)^2 + 4f^2 - 3f\right\} dx \qquad (a)$$

and has the value $f = 1$ at $x = 0$.

The Euler-Lagrange condition gives

$$-\frac{d}{dx}\left(2\frac{df}{dx}\right) + 8f - 3 = 0 \qquad (b)$$

or

$$\frac{d^2f}{dx^2} - 4f = \frac{3}{2} \qquad (c)$$

the solution of which is

$$f = A \sin 2x + B \cos 2x - \frac{3}{8} \qquad (d)$$

The specified boundary condition is $f(o) = 1$, whence

$$B = \frac{11}{8} \qquad (e)$$

Since $f(1)$ was not specified the function $f(x)$ must satisfy the natural boundary condition

$$\frac{\delta I}{\delta f_x}\bigg|_{x=1} = 0 = 2\frac{\delta f}{\delta x}\bigg|_{x=1} \qquad (f)$$

whence

$$\frac{df}{dx}\bigg|_{x=1} = 2A \cos 2 - \frac{11}{4} \sin 2 = 0 \qquad (g)$$

$$\therefore A = \frac{11}{8} \tan 2$$

giving,

$$f = \frac{1}{8}(11 \tan 2 \sin 2x + 11 \cos 2x - 3) \qquad (h)$$

If no restrictions had been placed upon the value of $\eta(x)$ at the ends of the interval we can see that $\eta(x)$ would be arbitrary, subject only to its being continuous in the interval. By choosing first the set of functions $\eta(x)$ such that $\eta(x) = 0$ at $x = x_1, x_2$ we would establish the Euler-Lagrange condition, by choosing the set of functions $\eta(x)$ such that $\eta(x_1) = 0$ we would establish the natural

boundary conditions at $x = x_2$, and by choosing the set of functions such that $\eta(x_2) = 0$ we would establish similar boundary conditions at $x = x_1$.

Example 1.6 Find the function $f(y)$ that extremizes the functional of example 1.5, if no boundary conditions at x_1 x_2 are given.

From the Euler-Lagrange condition we find, as before

$$f = A \sin 2x + B \cos 2x - \frac{3}{8} \qquad (a)$$

Since no restrictions have been placed upon $f(o)$ or $f(1)$ the extremizing function will satisfy the natural boundary conditions at $x = 0$ and $x = 1$, i.e.

$$\frac{df}{dx} = 0 \quad \text{at } x = 0,\ 1 \qquad (b)$$

These conditions give $A = 0$, $B = 0$, whence

$$f = -\frac{3}{8} \qquad (c)$$

is the required function.

——— : ———

We can also investigate the other types of functionals when less restrictive conditions are imposed at the ends of the interval of integration. Thus for the functional

$$F(f) = \int_{x_1}^{x_2} I(f,\ f_x,\ f_{xx},\ x)\,dx \qquad (1.42)$$

we obtain the expression

$$\int_{x_1}^{x_2} \left\{\frac{\partial I}{\partial f} - \frac{d}{dx}\frac{\partial I}{\partial f_x} + \frac{d^2}{dx^2}\frac{\partial I}{\partial f_{xx}}\right\}\eta\,dx + \left|\left(\frac{\partial I}{\partial f_x} - \frac{d}{dx}\frac{\partial I}{\partial f_{xx}}\right)\eta\right|_{x_1}^{x_2} + \left|\frac{\partial I}{\partial f_{xx}}\eta_x\right|_{x_1}^{x_2} = 0 \qquad (1.43)$$

If no restriction, apart from the continuity of f and f_x, is placed on $f(x)$, and hence on $\eta(x)$, we can establish the Euler-Lagrange conditions by taking the set of functions $\eta(x)$ with $\eta(x_1) = \eta_x(x_1) = \eta(x_2) = \eta_x(x_2) = 0$. By choosing the set of functions $\eta(x)$ such that three out of the four conditions above are satisfied we establish the natural boundary conditions one at the time. They are,

$$\frac{\partial I}{\partial f_x} - \frac{d}{dx}\frac{\partial I}{\partial f_{xx}} = 0 \qquad x = x_1, \quad x = x_2$$

$$\frac{\partial I}{\partial f_{xx}} = 0 \qquad x = x_1, \quad x = x_2 \tag{1.44}$$

We can carry out a similar operation for functionals including higher order derivatives. We can see by inspection that the first set of natural boundary conditions can be obtained from the Euler-Lagrange condition by taking the derivatives with respect to x and reducing these derivatives by one order. The second set is obtained by reducing these by one more order and so on.

Example 1.7 Find the natural boundary conditions for the functional

$$F(f) = \int_{x_1}^{x_2} \left[\left(\frac{d^4 f}{dx^4}\right)^4 + A\left(\frac{d^2 f}{dx^2}\right)^2\right] dx \tag{a}$$

The Euler-Lagrange condition is

$$\frac{d^4}{dx^4}\left(\frac{\partial I}{\partial f_{x(4)}}\right) + \frac{d^2}{dx^2}\left(\frac{\partial I}{\partial f_{xx}}\right) = 0 \tag{b}$$

Whence the natural boundary conditions are

$$\frac{d^3}{dx^3}\frac{\partial I}{\partial f_{x(4)}} + \frac{d}{dx}\frac{\partial I}{\partial f_{xx}} = 0 \quad \text{or} \quad \frac{d^7 f}{dx^7} + A\frac{d^3 f}{dx^3} = 0, \text{ on } x = x_1, x_2 \tag{c}$$

also

$$\frac{d^2}{dx^2}\frac{\partial I}{\partial f_{x(4)}} + \frac{\partial I}{\partial f_{xx}} = 0 \quad \text{or} \quad \frac{d^6 f}{dx^6} + A\frac{d^2 f}{dx^2} = 0 \text{ , on } x = x_1,\ x_2 \qquad (d)$$

similarly

$$\frac{d^5 f}{dx^5} = 0 \text{ and } \frac{d^4 f}{dx^4} = 0 \quad \text{on } x = x_1,\ x_2 \qquad (e)$$

——— : ———

It will be noticed that all the natural boundary conditions of a problem involve derivatives of the same order as, or of higher order than, the highest order of the derivatives in the functionals. The boundary conditions which may be specified for any particular problem can then be split into two groups. If the functional contains nth order derivatives, then boundary conditions relating only to derivatives up to the order n-1 are called essential, and those relating to higher order derivatives supplementary. If we have some process for finding the stationary value of a functional by using trial functions then these trial functions must be admissible functions and hence satisfy any specified essential boundary condition. In physical problems other boundary conditions will in fact relate to the natural boundary conditions, and these need not be satisfied by the trial functions. When the functions giving the functional its stationary value is obtained, it will automatically satisfy these boundary conditions.

1.6 Functionals with Several Dependent Variables

The procedure used in the earlier sections of this chapter for establishing the conditions for a stationary value for a functional can easily be extended to the case of functionals with several dependent variables. We will only consider the simplest case where

the functional involves two dependent variables, $f(x)$ and $g(x)$ and their first derivatives,

$$F(f, g) = \int_{x_1}^{x_2} I(f, f_x, g, g_x, x)dx \qquad (1,45)$$

where the functions f and g are to have the values f_1, g_1, f_2, g_2 respectively at $x = x_1, x_2$. The functions passing through $(x_1\ f_1)$ and $(x_2\ f_2)$ can be put in the form

$$h(x) = f(x) + \epsilon\, \eta(x) \qquad (1.46)$$

with $\eta(x_1) = \eta(x_2) = 0$, as before. The functions passing through $(x_1\ g_1)$ and $(x_2\ g_2)$ can likewise be put in the form

$$k(x) = g(x) + \epsilon\, \zeta(x) \qquad (1.47)$$

with $\zeta(x_1) = \zeta(x_2) = 0$.

The fact that we have used the same parameter ϵ does not imply that a given g can lead to a given f, or vice versa, since η and ζ are arbitrary (subjected to satisfying the conditions at the end of the interval and their being continuous).

In the above we use $f, g(x)$ to denote the actual functions which give the stationary value to F. We thus see that for all the possible values of ϵ, the one that gives the minimum value to the $F(f,g)$ will be $\epsilon = 0$, i.e. as before at $\epsilon = 0$ we have

$$\left.\frac{dF(h,k)}{d\epsilon}\right|_{\epsilon=0} = 0 \qquad (1.48)$$

Now

$$F(h,k) = \int_{x_1}^{x_2} I(h, h_x, k, k_x, x)dx \qquad (1.49)$$

whence

$$\frac{dF(h,k)}{d\epsilon} = \int_{x_1}^{x_2} \frac{\partial I}{\partial h}\frac{dh}{d\epsilon} + \frac{\partial I}{\partial h_x}\frac{dh_x}{d\epsilon} + \frac{\partial I}{\partial k}\frac{dk}{d\epsilon} + \frac{\partial I}{\partial k_x}\frac{dk_x}{d\epsilon}\,)dx = 0 \qquad (1.50)$$

Putting this

$$\frac{dh}{d\epsilon} = \eta \;,\; \frac{\partial h_x}{\partial \epsilon} = \eta_x \;,\; \frac{\partial k}{\partial \epsilon} = \zeta \;,\; \frac{\partial k_x}{\partial \epsilon} = \zeta_x \qquad (1.51)$$

and setting $h(x) = f(x)$, $k(x) = g(x)$ since $\epsilon = 0$, we obtain

$$\int_{x_1}^{x_2} \{\frac{\partial I}{\partial f}\eta + \frac{\partial I}{\partial f_x}\eta_x + \frac{\partial I}{\partial g}\zeta + \frac{\partial I}{\partial g_x}\zeta_x\}dx = 0 \qquad (1.52)$$

Integrating by parts the second and fourth terms in this expression we obtain

$$\int_{x_1}^{x_2} \{(\frac{\partial I}{\partial f} - \frac{d}{dx}\frac{\partial I}{\partial f_x})\eta + (\frac{\partial I}{\partial g} - \frac{d}{dx}\frac{\partial I}{\partial g_x})\zeta\}dx + \left|\frac{\partial I}{\partial f_x}\eta\right|_{x_1}^{x_2} + \left|\frac{\partial I}{\partial g_x}\zeta\right|_{x_1}^{x_2} \qquad (1.53)$$

The limit terms are zero since $\eta(x)$, $\zeta(x)$ are zero at the limits and since η and ζ are arbitrary each of the boundary terms in the integral must be zero,

$$\frac{\partial I}{\partial f} - \frac{d}{dx}\frac{\partial I}{\partial f_x} = 0 \;,\; \frac{\partial I}{\partial g} - \frac{d}{dx}\frac{\partial I}{\partial g_x} = 0 \qquad (1.54)$$

Similar terms would be obtained if there were more dependent functions in the functional. The Euler Lagrange conditions for functional with higher order derivatives would apply for each of the dependent variables.

Finally the natural boundary conditions for functionals with two or more dependent variables can be established as in the case of only one dependent variable. They will be found to have exactly the same form.

Example 1.8 Find the condition that the functional

$$F(f,g) = \int_{x_1}^{x_2} \{(\frac{df}{dx})^2 + (\frac{dg}{dx})^2 + (\frac{df}{dx})(\frac{dg}{dx}) + 4f^2 + g^2\}dx \qquad (a)$$

shall be stationary. f and g being specified at $x = x_1$ and $x = x_2$.

Here,

$$I = f_x^2 + g_x^2 + f_x g_x + 4f^2 + g^2 \qquad (b)$$

whence

$$\frac{\partial I}{\partial f} = 8f\ ,\ \frac{\partial I}{\partial f_x} = 2f_x + g_x\ ,\ \frac{\partial I}{\partial g} = 2g\ ,\ \frac{\partial I}{\partial g_x} = 2g_x + f_x \qquad (c)$$

The Euler-Lagrange conditions thus give

$$2\frac{d^2f}{dx^2} + \frac{d^2g}{dx^2} - 8f = 0,\ \frac{d^2f}{dx^2} + 2\frac{d^2g}{dx^2} - 2g = 0 \qquad (d)$$

——— : ———

1.7 Functionals with two or more Independent Variables

Consider the functional,

$$F(f) = \iint_D I(f, f_x, f_y, x, y)dx\,dy \qquad (1.55)$$

where the integration is carried out over some region D, and ϕ has specified values $f_S(x,y)$ at all points of the boundary or the region s(x,y). The function f(x,y) is to be continuous (figure 1.4).

As before we denote by f the function which gives a stationary value to this functional. Then any function such that satisfies $f = \phi_S$ on the boundary can be put in the form

$$h(x,y) = f(x,y) + \epsilon\,\eta(x,y) \qquad (1.56)$$

We know that whatever function we choose for η, the functional F(f) will be stationary for $\epsilon = 0$, i.e.

Now

$$\frac{dF}{d\epsilon} = \iint_D \left\{ \frac{\partial I}{\partial h} \frac{dh}{d\epsilon} + \frac{\partial I}{\partial h_x} \frac{dh_x}{d\epsilon} + \frac{\partial I}{\partial h_y} \frac{dh_y}{d\epsilon} \right\} dx\, dy \tag{1.58}$$

Using $\frac{dh}{d\epsilon} = \eta$, $\frac{dh_x}{d\epsilon} = \eta_x$, $\frac{dh_y}{d\epsilon} = \eta_y$, and putting $\epsilon = 0$, so that

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

$$\iint \left\{ \frac{\partial I}{\partial f} \eta + \frac{\partial I}{\partial f_x} \eta_x + \frac{\partial I}{\partial f_y} \eta_y \right\} dxdy = 0 \tag{1.59}$$

Integrating by parts the second and third term we obtain (Green's theorem)

$$\iint_D \frac{\partial I}{\partial f_x} \eta_x \, dxdy = - \iint_D \frac{\partial}{\partial x} \left(\frac{\partial I}{\partial f_x}\right) \eta \, dxdy + \int_S \frac{\partial I}{\partial f_x} \eta \, dy$$

$$\iint_D \frac{\partial I}{\partial f_y} \eta_y \, dxdy = - \iint_D \frac{\partial}{\partial y} \left(\frac{\partial I}{\partial f_y}\right) \eta \, dxdy - \int_S \frac{\partial I}{\partial f_y} \eta \, dx \tag{1.60}$$

The equation (1.59) can now be written after changing variables x, y to s on the boundary as

$$\iint_D \left\{ \frac{\partial I}{\partial f} - \frac{\partial}{\partial x} \left(\frac{\partial I}{\partial f_x}\right) - \frac{\partial}{\partial y} \left(\frac{\partial I}{\partial f_y}\right) \right\} \eta \, dxdy + \int_S \left(\frac{\partial I}{\partial f_x} \frac{dy}{ds} - \frac{\partial I}{\partial f_y} \frac{dx}{ds} \right) \eta \, ds \tag{1.61}$$

The second term, that is the line integral around the boundary in (1.61) is zero since η is zero on the boundary. This leads to the conclusion that, since η is arbitrary,

$$\frac{\partial I}{\partial f} - \frac{\partial}{\partial x} \left(\frac{\partial I}{\partial f_x}\right) - \frac{\partial}{\partial y} \left(\frac{\partial I}{\partial f_y}\right) = 0 \tag{1.62}$$

Problems involving higher order derivatives can be treated in a like manner. Thus for functionals of the form, for instance,

$$F(w) = \iiint_V I(w, w_x, w_y, w_z, w_{xx}, w_{yx}, w_{yz}, w_{xz}, w_{yy}, w_{yz}, w_{zz})\,dxdtdz \tag{1.63}$$

the Euler condition is

$$\frac{\partial I}{\partial w} - \left(\frac{\partial}{\partial x}\frac{\partial I}{\partial w_x} + \frac{\partial}{\partial y}\frac{\partial I}{\partial w_y} + \frac{\partial}{\partial z}\frac{\partial I}{\partial w_z}\right) + \frac{\partial^2}{\partial x^2}\frac{\partial I}{\partial w_{xx}} + \frac{\partial^2}{\partial x \partial y}\frac{\partial I}{\partial w_{xy}} + \frac{\partial^2}{\partial x \partial z}\frac{\partial I}{\partial w_{xz}}$$

$$+ \frac{\partial^2}{\partial y^2}\frac{\partial I}{\partial w_{yy}} + \frac{\partial^2}{\partial x \partial z}\frac{\partial I}{\partial w_{yz}} + \frac{\partial^2}{\partial z^2}\frac{\partial I}{\partial w_{zz}} = 0$$

——— : ———

Example 1.9 Find the condition for stationary of

$$F(w) = \iint_D \left\{\frac{B}{2}\left[\left(\frac{\partial^2 w}{\partial x^2}\right)^2 + \left(\frac{\partial^2 w}{\partial y^2}\right)^2 + 2\mu\left(\frac{\partial^2 w}{\partial x^2}\right)\left(\frac{\partial^2 w}{\partial y^2}\right) + 2(1-\mu)\left(\frac{\partial^2 w}{\partial x \partial y}\right)^2\right]\right.$$

$$\left. - pw \right\} dxdy \tag{a}$$

Here,

$$I = \frac{B}{2}\left[w_{xx}^2 + w_{yy}^2 + 2\mu\, w_{xx}\, w_{yy} + 2(1-\mu)\, w_{xy}^2\right] - pw \tag{b}$$

so that,

$$\frac{\partial I}{\partial w} = -p\ ,\quad \frac{\partial I}{\partial w_{xx}} = B(w_{xx} + \mu\, w_{yy}),\quad \frac{\partial I}{\partial w_{yy}} = B(w_{yy} + \mu\, w_{xx})$$

$$\frac{\partial}{\partial w_{xy}} = 2B(1-\mu)\, w_{xy} \tag{c}$$

Whence

$$-p + B\left\{\frac{\partial^2}{\partial x^2}(w_{xx} + \mu\, w_{yy}) + \frac{\partial^2}{\partial y^2}(w_{yy} + \mu\, w_{xx}) + 2(1-\mu)\frac{\partial^2}{\partial x \partial y}(w_{xy})\right\} = 0 \tag{d}$$

$$\frac{\partial^4 w}{\partial x^4} + 2\frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} - \frac{p}{B} = 0 \qquad (e)$$

——— : ———

Example 1.10 Consider the functional

$$F(f) = \frac{1}{2}\iint \{(\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2\}\, dxdy \qquad (a)$$

Here

$$I = \frac{1}{2}\{f_x^2 + f_y^2\} \qquad (b)$$

thus,

$$\frac{\partial I}{\partial f} = 0 \quad , \quad \frac{\partial I}{\partial f_x} = f_x \quad , \quad \frac{\partial I}{\partial f_y} = f_y \qquad (c)$$

The Euler-Lagrange condition is,

$$\frac{\partial}{\partial x}(f_x) + \frac{\partial}{\partial y}(f_y) = 0 \qquad (d)$$

which is the Laplace equation,

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0 \qquad (e)$$

Note that for this case the boundary term (equation (1.61)) is

$$\int_S (\frac{\partial I}{\partial f_x}\frac{dy}{ds} - \frac{\partial I}{\partial f_y}\frac{dx}{ds})\,\eta\, dS = \int_S (f_x\frac{dy}{ds} - f_y\frac{dx}{ds})\,\eta\, dS \qquad (f)$$

This gives (figure 1.5) the following natural boundary condition term,

$$\int_S \frac{\partial f}{\partial n}\,\eta\, dS \qquad (g)$$

1.8 Functionals with several Dependent and Independent Variables.

Let us now consider the case of a functional which depends on two separate functions f and g and applies on a domain D, function of x and y.

$$F(f, g) = \iint_D I(f, f_x, f_y, g, g_x, g_y, x, y)\, dxdy \qquad (1.65)$$

We can consider a f and g function as in 1.6 and obtain for the stationary condition

$$\frac{dF}{d\epsilon} = \iint_D \left\{ \frac{\partial I}{\partial h}\frac{dh}{d\epsilon} + \frac{\partial I}{\partial h_x}\frac{dh_x}{d\epsilon} + \frac{\partial I}{\partial h_y}\frac{dh_y}{d\epsilon} + \frac{\partial I}{\partial k}\frac{dk}{d\epsilon} + \frac{\partial I}{\partial k_x}\frac{dk_x}{d\epsilon} + \frac{\partial I}{\partial k_y}\frac{dk_y}{d\epsilon} \right\} dxdy \qquad (1.66)$$

which gives for $\epsilon = 0$

$$\left.\frac{dF}{d\epsilon}\right|_{\epsilon=0} = \iint_D \left\{ \frac{\partial I}{\partial f}\eta + \frac{\partial I}{\partial f_x}\eta_x + \frac{\partial I}{\partial f_y}\eta_y + \frac{\partial I}{\partial g}\zeta + \frac{\partial I}{\partial g_x}\zeta_x + \frac{\partial I}{\partial g_y}\zeta_y \right\} dxdy \qquad (1.67)$$

Integrating by parts we obtain,

$$\iint_D \left[\left\{ \frac{\partial I}{\partial f} - \frac{\partial}{\partial x}\left(\frac{\partial I}{\partial f_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial I}{\partial f_y}\right) \right\}\eta + \left\{ \frac{\partial I}{\partial g} - \frac{\partial}{\partial x}\left(\frac{\partial I}{\partial g_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial I}{\partial g_y}\right) \right\}\zeta \right] dxdy$$

$$+ \int_S \left[\left(\frac{\partial I}{\partial g_x}\frac{dy}{ds} + \frac{\partial I}{\partial f_y}\frac{dx}{ds}\right)\eta + \left(\frac{\partial I}{\partial g_x}\frac{dy}{ds} + \frac{\partial I}{\partial g_y}\frac{dx}{ds}\right)\zeta \right] dS \qquad (1.68)$$

As η and ζ are zero on S boundary we have the following two Euler equations

$$\frac{\partial I}{\partial f} - \frac{\partial}{\partial x}\left(\frac{\partial I}{\partial f_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial I}{\partial f_y}\right) = 0 \qquad (1.69)$$

$$\frac{\partial I}{\partial g} - \frac{\partial}{\partial x}\left(\frac{\partial I}{\partial g_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial I}{\partial g_y}\right) = 0$$

Example 1.11 Let us consider the following two dimensional functional, which applies for the internal energy of a plane stress structure,

$$F(u,v) = \frac{Eh}{2(1-\mu)^2} \iint_D \left\{ \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + 2\mu \left(\frac{\partial u}{\partial x}\right)\left(\frac{\partial v}{\partial y}\right) + \frac{1-\mu}{2}\left[\left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + 2 \frac{\partial u}{\partial y}\frac{\partial v}{\partial x} \right] \right\} dxdy \qquad (a)$$

where h is the thickness of the plate, E the modulus of elasticity, μ the Poisson's ratio and u, v the displacements in x, y directions.

We have

$$\frac{\partial I}{\partial u} = 0 , \quad \frac{\partial}{\partial x}\left(\frac{\partial I}{\partial u_x}\right) = \frac{Eh}{2(1-\mu^2)} (2u_{xx} + 2\mu\, v_{yx}),$$

$$\frac{\partial}{\partial y}\left(\frac{\partial I}{\partial u_y}\right) = \frac{Eh}{2(1-\mu^2)} \left[(1-\mu)(u_{yy} + v_{xy}) \right] \qquad (b)$$

$$\frac{\partial I}{\partial v} = 0 , \quad \frac{\partial}{\partial x}\left(\frac{\partial I}{\partial v_x}\right) = \frac{Eh}{2(1-\mu^2)} \left[(1-\mu)(v_{xx} + u_{yx}) \right]$$

$$\frac{\partial}{\partial y}\left(\frac{\partial I}{\partial v_y}\right) = \frac{Eh}{2(1-\mu^2)} (2v_{yy} + 2\mu\, u_{xy})$$

Which gives the following Euler's equations,

$$\frac{Eh}{1-\mu^2} \left\{ u_{xx} + \mu\, v_{yx} + \frac{(1-\mu)}{2} (u_{yy} + v_{xy}) \right\} = 0 \qquad (c)$$

$$\frac{Eh}{1-\mu^2} \left\{ v_{xx} + \mu\, u_{xy} + \frac{(1-\mu)}{2} (v_{xx} + u_{yx}) \right\} = 0$$

These are the equilibrium equations for a two dimensional plane stress elastic solid written in terms of displacements. In terms of stresses they are,

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau}{\partial y} = 0 \qquad \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau}{\partial x} = 0$$

where

$$\sigma_x = \frac{Eh}{1-\mu^2}(u_x + \mu v_y) \quad , \quad \sigma_y = \frac{Eh}{1-\mu^2}(v_y + \mu u_x), \quad \tau = \frac{E}{2(1+\mu)}(u_y + v_x)$$

——— : ———

1.9 The Variational Notation

We will now define the concept of a 'variation' in order to simplify the notation we have been using. Consider the case of the simple functional,

$$F(f) = \int_{x_1}^{x_2} I(f, f_x, x)\, dx \tag{1.70}$$

with an extremum for f.

The new function $h = f + \epsilon \eta$ will be written as $h = f + \delta f$ where $\delta \phi$ is called the 'variation' of f ($\delta f = \epsilon \eta$). Thus

$$h = f + \delta f$$

and

$$h_x = f_x + \delta f_x \tag{1.71}$$

as $\frac{d}{dx}(\delta f) = \frac{d}{dx}(\epsilon \eta(x)) = \delta \frac{d\eta}{dx} = \delta\left(\frac{df}{dx}\right)$

The quantities δf, δf_x are arbitrary in the interval $x_1 < x < x_2$.

The functional F can be expanded in the vicinity of the extremum solution f in function of ϵ.

$$F(h) = F(f) + \left.\frac{dF}{d\epsilon}\right|_{\epsilon=0} \epsilon + \frac{1}{2!}\left.\frac{d^2F}{d\epsilon^2}\right|_{\epsilon=0} \epsilon^2 + \ldots \tag{1.72}$$

The total increment of F function is

$$\Delta F = \frac{dF}{d\epsilon}_{\epsilon=0} \epsilon + \frac{1}{2!} \frac{d^2F}{d\epsilon^2}_{\epsilon=0} \epsilon^2 + \ldots \qquad (1.73)$$

The first term on the right hand side is defined as the first order increment or first variation of F; the second as the second order increment or second variation, etc. They are written as

$$\Delta F = \delta F + \frac{1}{2!} \delta^2 F + \ldots \qquad (1.74)$$

The first order increment can be written (equation 1.12, with $\epsilon = 0$) as

$$\delta F = \frac{dF}{d\epsilon}_{\epsilon=0} \epsilon = \epsilon \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial f} \eta + \frac{\partial I}{\partial f_x} \eta_x\right) dx = \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial f} \delta f + \frac{\partial I}{\partial f_x} \delta f_x\right) dx$$

as x does not vary. Finally,

$$\delta F = \frac{\partial F}{\partial f} \delta f + \frac{\partial I}{\partial f_x} \delta f_x \qquad (1.75)$$

Equation (1.75) shows that a variation can be applied to a functional in the same form as the differential of calculus, once the dependent variables are identified.

Although the variational notation and the notation used previously are equivalent, the former is easier to use. For instance, to obtain the Euler equation for the above functional we can do,

$$\delta F = \int_{x_1}^{x_2} \delta I \, dx = \int_{x_1}^{x_2} \left(\frac{\partial I}{\partial f} \delta f + \frac{\partial I}{\partial f_x} \delta f_x\right) dx \qquad (1.76)$$

which integrating by parts, gives

$$\delta F = \int_{x_1}^{x_2} \left[\frac{\partial I}{\partial f} - \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) \right] \delta f \, dx - \left[\delta f \, \frac{\partial I}{\partial f_x} \right]_{x_1}^{x_2} = 0 \qquad (1.77)$$

Thus the Euler condition is

$$\frac{\partial I}{\partial f} - \frac{d}{dx} \left(\frac{\partial I}{\partial f_x} \right) = 0 \qquad (1.78)$$

Let us now consider the case of a function of two variables (equation 1.45)

$$F(f, g) = \int_{x_1}^{x_2} I(f, g, f_x, g_x, x) dx \qquad (1.79)$$

with the new functions,

$$\begin{aligned} h(x) &= f(x) + \epsilon \eta(x) \quad \text{and} \quad h_x = f_x + \epsilon \eta_x \\ k(x) &= g(x) + \epsilon \zeta(x) \quad \text{and} \quad k_x = g_x + \epsilon \zeta_x \end{aligned} \qquad (1.80)$$

We can expend F in the proximity of f, g solution, assuming x is not varied and the limits are fixed. Thus,

$$F(h,k) = F(f, g) + \epsilon \left. \frac{dF}{d\epsilon} \right|_{\epsilon=0} + \frac{1}{2!} \left. \frac{d^2F}{d\epsilon^2} \right|_{\epsilon=0} + \ldots \qquad (1.81)$$

or $$\Delta F = \delta F + \frac{1}{2!} \delta^2 F + \ldots$$

The first order increment is (equation 1.52)

$$\begin{aligned} \delta F &= \left. \frac{dF}{d\epsilon} \right|_{\epsilon=0} \epsilon = \frac{\partial F}{\partial f} \eta \epsilon + \frac{\partial F}{\partial f_x} \eta_x \epsilon + \frac{\partial F}{\partial g} \zeta \epsilon + \frac{\partial F}{\partial g_x} \zeta_x \epsilon = \\ &= \frac{\partial F}{\partial f} \delta f + \frac{\partial F}{\partial f_x} \delta f_x + \frac{\partial F}{\partial g} \delta g + \frac{\partial F}{\partial g_x} \delta g_x = 0 \end{aligned} \qquad (1.82)$$

Integrating by parts we will obtain the two Euler equations

$$\delta F = \int_{x_1}^{x_2} \left\{ \left(\frac{\partial I}{\partial f} - \frac{d}{dx} \frac{\partial I}{\partial f_x} \right) \delta f + \left(\frac{\partial I}{\partial g} - \frac{d}{dx} \left(\frac{\partial I}{\partial g_x} \right) \delta g \right\} dx + \left| \frac{\partial I}{\partial f_x} \delta f \right|_{x_1}^{x_2} + \left| \frac{\partial I}{\partial g_x} \delta g \right|_{x_1}^{x_2} \tag{1.83}$$

It is interesting to generalize the variational notation to a function of 'n' variables, such as

$$F(f^1, f^2, f^3 \ldots) \tag{1.84}$$

The increment of this functional is now defined as

$$\Delta F = \delta F + \frac{1}{2!} \delta^2 F + \ldots \tag{1.85}$$

where

$$\delta F = \sum_i \frac{\partial F}{\partial f^i} \delta f^i \quad ; \quad \delta^2 F = \delta(\delta F) = \sum_j \sum_i \frac{\partial^2 F}{\partial f^i \partial f^j} \delta f^i \delta f^j$$

The second increment for a functional like $\int I(f^1, f^2, x)dx$ is,

$$\frac{d}{d\epsilon} \left(\frac{dF}{d\epsilon} \right) = \frac{d}{d\epsilon} \left(\frac{\partial F}{\partial h_1} \frac{\partial h_1}{\partial \epsilon} + \frac{\partial F}{\partial h_2} \frac{\partial h_2}{\partial \epsilon} \right) =$$

$$= \frac{\partial^2 F}{\partial h_1^2} \left(\frac{\partial h_1}{\partial \epsilon} \right)^2 + 2 \frac{\partial^2 F}{\partial h_1 \partial h_2} \left(\frac{\partial h_1}{\partial \epsilon} \right) \left(\frac{\partial h_2}{\partial \epsilon} \right) + \frac{\partial^2 F}{\partial h_2^2} \left(\frac{\partial h_2}{\partial \epsilon} \right)^2 \tag{1.86}$$

When $\epsilon = 0$ we have,

$$\frac{d}{d\epsilon} \left(\frac{dF}{d\epsilon} \right)_{\epsilon=0} = \frac{\partial^2 F}{\partial f_1^2} \eta_1^2 + 2 \frac{\partial^2 F}{\partial f_1 \partial f_2} \eta_1 \eta_2 + \frac{\partial^2 F}{\partial f_2^2} \eta_2^2 \tag{1.87}$$

Thus

$$\delta^2 F = \frac{\partial^2 F}{\partial f_1^2} \delta f_1^2 + 2 \frac{\partial^2 F}{\partial f_1 \partial f_2} \delta f_1 \delta f_2 + \frac{\partial^2 F}{\partial f_2^2} \delta f_2^2 \tag{1.88}$$

Example 1.12 Using the variational notation, deduce the Euler equation of the following functional,

$$F = \iint \left\{ \left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2 - 2fp \right\} dxdy \qquad (a)$$

which can be written

$$F = F(f_{xx}, f_{xy}, f_{yy}, f, x, y). \qquad (b)$$

Thus,

$$\delta F = \frac{\partial F}{\partial f_{xx}} \delta f_{xx} + \frac{\partial F}{\partial f_{xy}} \delta f_{xy} + \frac{\partial F}{\partial f_{yy}} \delta f_{yy} + \frac{\partial F}{\partial} \delta f \qquad (c)$$

$$= \iint \left\{ 2f_{xx}\delta f_{xx} + 4f_{xy}\delta f_{xy} + 2f_{yy}\delta f_{yy} - 2p\delta f \right\} dxdy \qquad (d)$$

Integrating this last equation by parts and neglecting the boundary terms, which are going to be satisfied, we obtain

$$\delta F = \iint \left\{ 2\frac{\partial^4 f}{\partial x^4} + 4\frac{\partial^2 f}{\partial x^2 \partial y^2} + 2\frac{\partial^4 f}{\partial y^4} - 2p \right\} \delta f \, dxdy \qquad (e)$$

The Euler Lagrange equation is,

$$\frac{\partial^4 f}{\partial x^4} + 2\frac{\partial^4 f}{\partial x^2 \partial y^2} + \frac{\partial^4 f}{\partial y^4} = \frac{p}{2} \qquad (f)$$

——— : ———

1.10 Subsidiary Conditions

In certain cases we want the variations to satisfy, in addition to the boundary conditions, certain other conditions, called subsidiary. These conditions can be introduced using Lagrange multipliers.

Let us review briefly what they are before proposing their use in

TABLE 1.1

Some Functionals and their Corresponding Euler-Lagrange Equations

FUNCTIONAL	EULER-LAGRANGE EQUATIONS	BOUNDARY TERMS
$\int_{x_1}^{x_2} I(x,f,f_x)dx + h_2(x,f)\Big\|_{x_2} - h_1(x,f)\Big\|_{x_1}$	$\frac{\partial I}{\partial f} - \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right) = 0$	$\left(\frac{\partial I}{\partial f_x} - \frac{\partial H_i}{\partial f}\right)\delta f\Big\|_{x=x_i} = 0$ $\quad i=1,2.$
$\int_{x_1}^{x_2} I(x,f,f_x,f_{xx})dx + h_2(x,f,f_x)\Big\|_{x_2} - H_1(x,f,f_x)\Big\|_{x_1}$	$\frac{\partial I}{\partial f} - \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right) + \frac{d^2}{dx^2}\left(\frac{\partial I}{\partial f_{xx}}\right) = 0$	$\left(\frac{\partial I}{\partial f_{xx}} + \frac{\partial H_i}{\partial f_x}\right)\delta f_x\Big\|_{x=x_i} = 0$ $\left(\frac{\partial I}{\partial f_x} - \frac{d}{dx}\frac{\partial I}{\partial f_{xx}} + \frac{\partial H_i}{\partial f}\right)\delta f\Big\|_{x=x_i} = 0$ $i=1,2.$
$\int_{x_1}^{x_2} I(x,f,f_x,f_{xx},\ldots f_{x(n)})dx$	$\frac{\partial I}{\partial f} - \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right) - \frac{d}{dx^2}\left(\frac{\partial I}{\partial f_{xx}}\right) + \ldots + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial F}{\partial x^n}\right) = 0$	
$\iint_D I(x,y,f,f_x,f_y)dxdy + \int \left[H_1(x,y,f)n_2 - H_2(x,y,f)n_1\right]dS$	$\frac{\partial I}{\partial f} - \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right) - \frac{d}{dy}\left(\frac{\partial I}{\partial f_y}\right) = 0$	$\left[\left(\frac{\partial I}{\partial f_x} + \frac{\partial H_1}{\partial f}\right)n_2 - \left(\frac{\partial I}{\partial f_y} + \frac{\partial H_2}{\partial f}\right)n_1\right]\delta f = 0$ n_1 n_2 are direction cosines of normal to boundary S with respect to x and y.

TABLE 1.1 (contd)

FUNCTIONAL	EULER-LAGRANGE EQUATIONS	BOUNDARY TERMS
$\iint I(x,y,f,g,f_x,f_y,g_x,g_y)\,dA$	$\frac{\partial I}{\partial f} - \frac{d}{dx}\left(\frac{\partial I}{\partial f_x}\right) - \frac{d}{dy}\left(\frac{\partial I}{\partial f_y}\right) = 0$	$\frac{\partial I}{\partial f_x}\,\delta f\Big\|_{x_i} = 0 \quad , \quad \frac{\partial I}{\partial g_x}\,\delta g\Big\|_{x_i} = 0$
	$\frac{\partial I}{\partial g} - \frac{d}{dx}\left(\frac{\partial I}{\partial g_x}\right) - \frac{d}{dy}\left(\frac{\partial I}{\partial g_y}\right) = 0$	$\frac{\partial I}{\partial f_y}\,\delta f\Big\|_{y_i} = 0 \quad , \quad \frac{\partial I}{\partial g_y}\,\delta g\Big\|_{x_i} = 0$

functionals. Consider a function $f(x,y,z)$ of which we want to obtain the stationary value

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dx = 0 \tag{1.89}$$

subject to the two constraints

$$\begin{aligned} g_1(x\ y\ z) &= 0 \\ g_2(x\ y\ z) &= 0 \end{aligned} \tag{1.90}$$

Note that now we will only have one independent variable.

We can differentiate (1.90)

$$\begin{aligned} dg_1 &= \frac{\partial g_1}{\partial x} dx + \frac{\partial g_1}{\partial y} dy + \frac{\partial g_1}{\partial z} dz = 0 \\ dg_2 &= \frac{\partial g_2}{\partial x} dx + \frac{\partial g_2}{\partial y} dy + \frac{\partial g_2}{\partial z} dz = 0 \end{aligned} \tag{1 91}$$

Let us multiply (1.91) by the unknown parameters $\lambda_1\ \lambda_2$ and add to (1.89). Then, we obtain

$$\begin{aligned} &(\frac{\partial f}{\partial x} + \lambda_1 \frac{\partial g_1}{\partial x} + \lambda_2 \frac{\partial g_2}{\partial x})dx + (\frac{\partial f}{\partial y} + \lambda_1 \frac{\partial g_1}{\partial y} + \lambda_2 \frac{\partial g_2}{\partial y})dy \\ &\quad + (\frac{\partial f}{\partial z} + \lambda_1 \frac{\partial g_1}{\partial z} + \lambda_2 \frac{\partial g_2}{\partial z})dz = 0 \end{aligned} \tag{1.92}$$

This gives 3 equations which add to (1.90) permitts to determine the five unknowns x, y, z, $\lambda_1\ \lambda_2$ - The parameters $\lambda_1\ \lambda_2$ are known as Lagrange multipliers, as sometimes they can be given a physical meaning.

Finally, we can now write the problem as the minimization of a new functional

$$F = f + \lambda_1 g_1 + \lambda_2 g_2 \tag{1.93}$$

We can miminize (1.93) with repsect to x, y, z, λ_1 and λ_2 and obtain

$$\frac{\partial f}{\partial x} + \lambda_1 \frac{\partial g_1}{\partial x} + \lambda_2 \frac{\partial g_2}{\partial x} = 0$$

$$\frac{\partial f}{\partial y} + \lambda_1 \frac{\partial g_1}{\partial y} + \lambda_2 \frac{\partial g_2}{\partial y} = 0 \qquad (1.94)$$

$$\frac{\partial f}{\partial z} + \lambda_1 \frac{\partial g_1}{\partial z} + \lambda_2 \frac{\partial g_2}{\partial z} = 0$$

plus $g_1 = 0$, $g_2 = 0$

Example 1.14 Let us find the extreme value of a f(x,y) function.

$$f(x,y) = x^2 + y^2 + 2 \qquad (a)$$

subjected to the constraint

$$g(x,y) = x + y - 1 = 0 \qquad (b)$$

The Lagrangian multiplier λ allows to form a new function

$$F(x,y) = f(x,y) + \lambda g(x,y) \qquad (c)$$

$$= x^2 + y^2 + 2 + \lambda(x + y - 1)$$

This function ca now be extremized with respect to x, y, λ. It gives

$$\frac{\partial F}{\partial x} = 2x + \lambda = 0$$

$$\frac{\partial F}{\partial y} = 2y + \lambda = 0 \qquad (d)$$

$$\frac{\partial F}{\partial \lambda} = x + y - 1 = 0$$

The solution of (d) gives,

$$x = \tfrac{1}{2} , \quad y = \tfrac{1}{2} , \quad \lambda = -1 \qquad (e)$$

Thus the extremum of f(x,y) under the subsidiary condition (b), is

$$f(x,y) = (\tfrac{1}{2})^2 + (\tfrac{1}{2})^2 + 2 = \frac{5}{2} \qquad (f)$$

——— : ———

Let us now assume to have a functional,

$$F = \int_{x_1}^{x_2} I(x, f, f_x)dx \tag{1.95}$$

subjected to the subsidiary conditions,

$$J = \int_{x_1}^{x_2} G(x, f, f_x)dx = 0 \tag{1.96}$$

Using the Lagrange multiplier λ we can write a new functional

$$F + \lambda J \tag{1.97}$$

We can minimize (1.97) with respect to f, f_x and λ

$$\delta(F + \lambda J) = 0 \tag{1.98}$$

$$\frac{\partial(F+\lambda J)}{\partial f}\delta f + \frac{\partial(F+\lambda J)}{\partial f_x}\delta f_x + \frac{\partial(F+\lambda J)}{\partial \lambda}\delta\lambda = 0$$

The first two terms give an Euler equation

$$\frac{d}{dx}\frac{\partial}{\partial f_x}(I + \lambda G) - \frac{\partial}{\partial f}(I + \lambda G) = 0 \tag{1.99}$$

and the third, the condition

$$J = 0 \tag{1.100}$$

If the functional depends on two variables f, g, we have

$$F = \int_{x_1}^{x_2} I(x, f, g, f_x, g_x)dx \tag{1.101}$$

plus a subsidiary condition

$$J = \int_{x_1}^{x_2} G(x, f, g, f_x, g_x)dx = 0 \tag{1.102}$$

The new functional is,

$$F + \lambda J \tag{1.103}$$

Equation (1.103) should satisfy

$$\delta(F + \lambda J) = 0 \tag{1.104}$$

which gives the following Euler equations

$$\frac{d}{dx}\frac{\partial}{\partial f_x}(I + \lambda G) - \frac{\partial}{\partial f}(I + \lambda G) = 0 \tag{1,105}$$

$$\frac{d}{dx}\frac{\partial}{\partial g_x}(I + \lambda G) - \frac{\partial}{\partial g}(I + \lambda G) = 0$$

plus the subsidiary condition $J = 0$.

The same procedure is valid for problems with more variables.

Bibliography

Courant, R. and D. Hilbert, "Methods of Mathematical Physics". Interscience Publ., 1953.

Forray, M.J. "Variational Calculus in Science and Engineering". McGraw Hill, 1968.

Hildebrand, F,B., "Methods of Applied Mathematics". Prentice Hall 1952.

Kantorovich, L.V. and V.I. Krylov, "Approximate Methods of Higher Analysis". Noordhoff Ltd., 1958.

Lanczos, C., "The Variational Principle of Mechanics". University of Toronto Press, 1949.

Weinstock. R., "Calculus of Variations". McGraw-Hill Co., 1952.

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS

APLICACION DE LOS ELEMENTOS FINITOS A FLUJO POTENCIAL

DR. GERARDO HIRIART

MAYO, 1978.

G. HIRIART

Aplicación de Elementos Finitos a flujo potencial

Durante este curso hemos visto y veremos muchas soluciones analíticas a problemas prácticos. (Estas soluciones son exactas en cuanto a las matemáticas, las únicas aproximaciones están en suponer que el fluído no tiene viscosidad).

La medula en todo problema de flujo con potencial está en resolver la ecuación de Laplace. LO QUE HACE DIFÍCIL SU SOLUCIÓN, (o imposible) SON LAS CONDICIONES DE FRONTERA. Mediante los métodos que tratamos con variable compleja, nos limitamos a casos bidimensionales, además sólo podremos resolver una cantidad limitada (analíticamente) de problemas prácticos.

A continuación veremos un método que es matemáticamente aproximado, aplicado a un caso ya aproximado al suponer flujo no viscoso, pero SUMAMENTE práctico para el ingeniero que generalmente busca soluciones realistas para problemas imposibles.

Este método, el elemento finito, creado por ingenieros hace pocos años dando excelentes resultados fue duramente criticado por los matemáticos quienes lo volvieron a reinventar pero con todos los pasos matemáticos probados rigurosamente y ampliados a casos mas generales.

El objetivo de este apunte es enseñar al alumno a profitar de tan útil herramienta en un solo caso determinado, el de la ecuación de Laplace, olvidándonos de todo el respaldo matemático y mostrando solo lo esencial que nos lleve a resultados prácticos, tratando por supuesto de no caer en la inmoralidad matemática. en beneficio de lo simple.
~~(poesía por favor no objetivos)~~.

En general el método del elemento finito consiste en:

1.- Dividir el área donde se resolverá la ecuación de Laplace $\nabla^2\phi = 0$ en varios elementos finitos. (pueden ser triángulos, cuadrados, rectangulos o incluso elementos de lados curvos.)

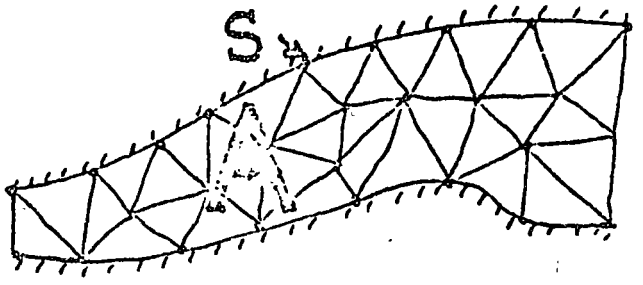
2.- Asumir que el potencial ϕ varía, dentro de cada elemento, en forma lineal, cuadrática o cúbica etc... (dependiendo de la exactitud que necesitemos)

3.- Aplicar algún método para resolver la ecuación de Laplace dentro de cada elemento

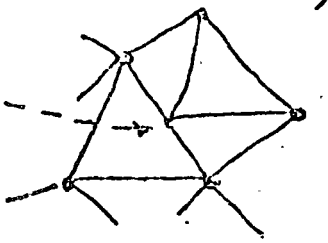
4.- Finalmente encontrar el valor de ϕ en cada nudo de los elementos, lo que será el resultado final del problema.

Para el caso que estudiaremos consideremos lo siguiente:

1.- Para el area A y contorno S mostrado en la figura, dividiremos el plano en triangulos (los que pueden ser de forma y tamaño arbitrarios)

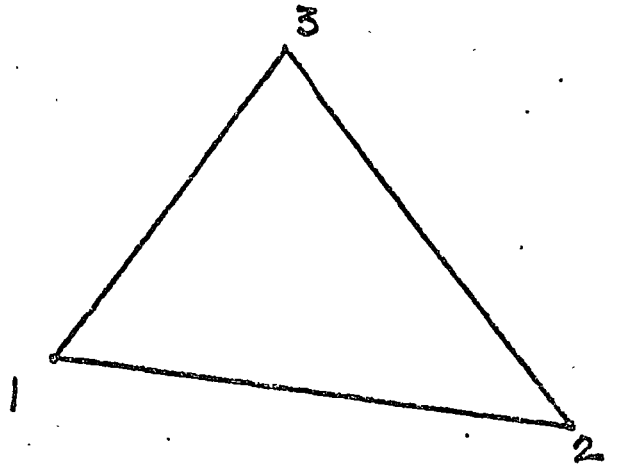


(un detalle importante que se hará obvio mas adelante es que los triangulos deben quedar todos unidos por sus vértices, y nunca un caso como este



2.- Debemos escoger cómo haremos variar ϕ de un nudo a otro. ¿Lineal, cuadrático etc...? En general usaremos para la ecuación de Laplace una variación cuadrática (ya que hay segundas derivadas) pero en el caso que ilustraremos a continuación lo haremos en variación lineal para entender más fácilmente el método.

Tomemos un triángulo cualquiera de la figura anterior y llamemos a los nodos 1, 2, 3.



En cada nudo, ϕ tendrá un valor constante de un nodo a otro, ϕ_1, ϕ_2 y ϕ_3. Como ϕ varía linealmente entonces tendremos

$$\phi = \{ \phi_1 [(x_2 y_3 - x_3 y_2) + (y_2 - y_3) x + (x_3 - x_2) y]$$

$$+ \phi_2 [(x_3 y_1 - x_1 y_3) + (y_3 - y_1) x + (x_1 - x_3) y]$$

$$+ \phi_3 [(x_1 y_2 - x_2 y_1) + (y_1 - y_2) x + (x_2 - x_1) y] \} / 2 \cdot \text{Area}$$

o bien $\phi = N_i \phi_i$ para abreviar la notación.

[NOTA no se compliquen tanto tratando de probar que esto es correcto ya que más adelante usaremos otro tipo de coordenadas mucho más prácticas]

Es fácil imaginarse como se complican las cosas al usar una variación cúbica o cuadrática.
Es importante fijarse en que los valores ϕ_i son constantes. Haremos mención el problema cuando encontremos cuánto valen esas constantes.
Para familiarizarnos con la notación a usar, los valores ϕ_1, ϕ_2, ϕ_3 del elemento "e" los llamaremos ϕ_i^e y el valor (variable) de ϕ dentro del elemento "e" será ϕ^e (Es decir ϕ_i^e son constantes, ϕ^e son variables ~~f~~ función de la posición)

3.- Resolver la ecuación de Laplace en cada triángulo satisfaciendo las condiciones de frontera.
¿Pero cómo resolver esta ecuación diferencial dentro de un triangulo?: He ahí el corazón del método y su belleza matemática.
Hay muchos métodos para hacer ésto, todos con nombres de personajes importantes; Galerkin, Kantorovič, Rayleigh, Stodola, residuos y muchos otros.
En general podemos decir que todos buscan una solución aproximada para $\nabla^2\phi = 0$, de la siguiente forma.

a.- Tomemos una función aproximada (a ojo) que llamaremos $\tilde{\phi}$ para distinguirla de la solución exacta que es ϕ. Con $\tilde{\phi}$ la ecuación de Laplace no nos dará cero sino que un residuo R

$$\nabla^2\tilde{\phi} = R$$

mientras menor sea ese residuo, mejor será nuestra solución.

b.- Busquemos en los libros de matemáticas si existe algún método como minimizar ese residuo. Encontraremos muchos, pero aprovechando la experiencia de otros ingenieros escogemos el método de Galerkin (que está de moda y produce muy buenos resultados-prácticos).

c.- Método de Galerkin: Para el caso de la ecuación de Laplace que estamos estudiando, tomamos a ojo una solución $\tilde{\phi}$, la que nos producirá un residuo R

$$\nabla^2 \tilde{\phi} = R$$

Galerkin dice que para minimizar ese residuo R, hay que hacerlo ortogonal a la propia función $\tilde{\phi}$. (En otras palabras R debe estar fuera del espacio completo de las $\tilde{\phi}$), (si todavía no se le aclara bien qué quiere decir éste, conformémonos con la fórmula), es decir

$$\int_A R \cdot \tilde{\phi} \, dA = 0$$

obviamente $\tilde{\phi}=0$ o $\tilde{\phi}=C$ o $\tilde{\phi}=Ax+By$ será siempre una solución. Pero que no cumpla satisface las condiciones de frontera.

Reemplazando R por $\nabla^2\tilde{\phi}$ e incluyendo las condiciones naturales de frontera

$$\int_S \tilde{\phi} \frac{\partial \tilde{\phi}}{\partial n} \, ds + \int_A \tilde{\phi} \, \nabla^2 \tilde{\phi} \, dA = 0$$

, o bien

$$\int_S \tilde{\phi} \frac{\partial \tilde{\phi}}{\partial n} \, ds + \iint \tilde{\phi} \left[\frac{\partial^2 \tilde{\phi}}{\partial x^2} + \frac{\partial^2 \tilde{\phi}}{\partial y^2} \right] dx \, dy = 0$$

y ahora usando un poco de cálculo para la segunda de las integrales

$$= \int \left[\int \tilde{\phi} \frac{\partial^2 \tilde{\Phi}}{\partial x^2} dx \right] dy + \int \left[\int \tilde{\phi} \frac{\partial^2 \tilde{\Phi}}{\partial y^2} dy \right] dx$$

e integrando por partes (¿se acuerda de $\int u dv$ ... ?) tenemos.

$$= \int \left[\tilde{\Phi} \frac{\partial \tilde{\Phi}}{\partial x} - \cancel{\int} \frac{\partial \tilde{\Phi}}{dx} \frac{\partial \tilde{\Phi}}{\partial x} dx \right] dy + \int \left[\tilde{\phi} \frac{\partial \tilde{\Phi}}{\partial y} - \int \frac{\partial \tilde{\Phi}}{\partial y} \frac{\partial \tilde{\Phi}}{\partial y} dy \right] dx$$

$$= -\cancel{\int} \iint \left[\left(\frac{\partial \tilde{\Phi}}{\partial x} \right)^2 + \left(\frac{\partial \tilde{\Phi}}{\partial y} \right)^2 \right] dx dy + \int_s \tilde{\phi} \frac{\partial \tilde{\Phi}}{\partial x} dy + \int \tilde{\phi} \frac{\partial \tilde{\Phi}}{\partial y} dx$$

Las dos ultimas integrales producen $\int \tilde{\phi} \frac{\partial \tilde{\phi}}{\partial n} ds$

Ahora podemos definir una función de $\tilde{\Phi}$, $I(\tilde{\phi})$ así':

$$I(\tilde{\phi}) = \frac{1}{2} \rho \int_A \left[\left(\frac{\partial \tilde{\Phi}}{\partial x} \right)^2 + \left(\frac{\partial \tilde{\Phi}}{\partial y} \right)^2 \right] dA - \rho \int_S \tilde{\phi} \frac{\partial \tilde{\Phi}}{\partial n} ds$$

La solución mas correcta para $\tilde{\Phi}$ será aquella que haga mínima la función $I(\tilde{\Phi})$.

Si encontráramos la solución exacta Φ, entonces $I(\Phi) \equiv$ y la ecuación quedaría como

$$\frac{1}{2} \rho \int_A \left[\left(\frac{\partial \phi}{\partial x} \right)^2 + \left(\frac{\partial \phi}{\partial y} \right)^2 \right] dA = \rho \int_S \phi \frac{\partial \phi}{\partial n} ds$$

¡Obsérvese, ~~por favor~~, cuál es el significado físico de esta ecuación! A la izquierda está la energía cinética del fluído y a la derecha (aunque no tan obvio) el trabajo instantáneo hecho por el contorno S para poner el fluído en movimiento.

Vale la pena aquí resaltar algunos detalles.

Para el método del elemento finito es necesario transformar la ecuación diferencial en una ecuación integral. Esto lo hicimos por el método de Galerkin, también pudo haberse hecho como un planteamiento de Cálculo Variacional, (el que quiera profundizar en el método del elemento finito debe saber usar las herramientas y sofisticaciones del Cálculo variacional).

d.- Hasta aquí tenemos que para cada elemento

$$I(\tilde{\phi}^e) = \frac{1}{2}\rho \int_A \left[\left(\frac{\partial \tilde{\phi}^e}{\partial x}\right)^2 + \left(\frac{\partial \tilde{\phi}^e}{\partial y}\right)^2 \right] dA - \rho \int_S \tilde{\phi}^e \frac{\partial \tilde{\phi}}{\partial n} ds$$

¿Y cómo minimizamos la función $I(\phi)$?
muy fácil, como $\tilde{\phi}^e$ en cada elemento era

$$\tilde{\phi}^e = N_i^e \; \tilde{\phi}_i^e$$

(por si no recuerda, N_i^e era todo eso de $(x_2 y_3 - x_3 y_2) + (y_2 - y_3)x \ldots$ etc)
y el N_i^e es algo fijo del problema que nosotros ya escogimos al decir que la variación de $\tilde{\phi}^e$ era lineal o cuadrática en el elemento e, tenemos que encontrar esas constantes $\tilde{\phi}_i^e$ que hagan mínima la expresión $I(\tilde{\phi}^e)$

Por lo tanto nos convencimos que minimizar $I(\tilde{\phi}^e)$ en el elemento e, es

$$\frac{\partial I(\tilde{\phi}^e)}{\partial \tilde{\phi}_i^e} = 0$$

mediante este procedimiento (método de Ritz) encontraremos los $\tilde{\phi}_i^e$ que minimizan I. Ese será el resultado final.

Como $\varphi^e = N_i \varphi_i$

o lo que es lo mismo $= N_1^e \phi_1^e + N_2^e \phi_2^e + N_3^e \phi_3^e$

(para variación lineal)

Si lo reemplazamos en la Integral de $I^e(\phi^e)$ tenaremos.

$$I^e(\tilde{\phi}^e) = \frac{1}{2} \rho \int_A \left[\left(\frac{\partial N_i^e \tilde{\phi}_i^e}{\partial x} \right)^2 + \left(\frac{\partial N_i^e \tilde{\phi}_i^e}{\partial y} \right)^2 \right] dA - \rho \int_S N_i^e \tilde{\phi}_i^e \frac{\partial \tilde{\phi}^e}{\partial n} \, ds$$

(*) ya le explicaré por qué esto.

Como los $\tilde{\phi}_i^e$ son constantes

$$I^e(\tilde{\phi}^e) = \frac{1}{2} \rho \int_A \left[\left(\tilde{\phi}_1^e \frac{\partial N_1^e}{\partial x} + \tilde{\phi}_2^e \frac{\partial N_2^e}{\partial x} + \tilde{\phi}_3^e \frac{\partial N_3^e}{\partial x} \right)^2 + \left(\tilde{\phi}_1^e \frac{\partial N_1^e}{\partial y} + \tilde{\phi}_2^e \frac{\partial N_2^e}{\partial y} \right.\right.$$

$$\left.\left. + \tilde{\phi}_3^e \frac{\partial N_3^e}{\partial y} \right)^2 dA - \rho \int_S \left(\tilde{\phi}_1^e N_1^e + \tilde{\phi}_2^e N_2^e + \tilde{\phi}_3^e N_3^e \right) \frac{\partial \tilde{\phi}}{\partial n} \, ds$$

(*) ese término $\frac{\partial \tilde{\phi}^e}{\partial n}$ no lo reemplacé por que sobre la frontera S, casi siempre conocemos $\frac{\partial \phi}{\partial n}$ (por ejemplo si la frontera es una línea de corriente, $\frac{\partial \phi}{\partial n} = 0$).

Ahora minimizamos $I^e(\tilde{\phi}^e)$ con respecto a cada ϕ_i^e haciendo

$$\frac{\partial I^e(\tilde{\phi}^e)}{\partial \tilde{\phi}_i^e} = 0 \qquad \text{y nos queda}$$

$$0 = \rho \int_A \left[\left(\tilde{\phi}_i^e \frac{\partial N_i^e}{\partial x} \right) \frac{\partial N_j^e}{\partial x} + \left(\tilde{\phi}_i^e \frac{\partial N_i^e}{\partial y} \right) \frac{\partial N_j^e}{\partial y} \right] dA$$

$$- \rho \int_S N_j^e \frac{\partial \tilde{\phi}}{\partial n} \, dA$$

lo que escrito en forma mas compacta quedará:

$$\tilde{\phi}_i^e \int \left[\frac{\partial N_i^e}{\partial x} \frac{\partial N_j^e}{\partial x} + \frac{\partial N_i^e}{\partial y} \frac{\partial N_j^e}{\partial y} \right] dA = \int N_j^e \frac{\partial \tilde{\phi}}{\partial n} \, ds$$

Resumen.

1.- Escogimos dividir el area A en triángulos, poniendolos mas agrupados en aquellas regiones donde esperamos variaciones más grandes de velocidades

2.- Escogimos una variación cuadrática (o lineal) de ϕ dentro de cada triángulo. Así podremos escribir $\phi^e = N_i^e \phi_i^e$
donde N_i es una expresión que no cambiará de un problema otro (estamos hablando de la ecuación de Laplace) por lo tanto cuando obtengamos N_i para variación cuadrática (o lineal) de ϕ esa nos servirá para cualquier otro problema.

3.- Basado en el método de Galerkin obtuvimos una función $I^e(\tilde{\phi}^e)$ para cada elemento

$$I^e(\tilde{\phi}^e) = \frac{1}{2}\rho \int_A \left[\left(\tilde{\phi}_i^e \frac{\partial N_i^e}{\partial x} \right)^2 + \left(\tilde{\phi}_i^e \frac{\partial N_i^e}{\partial y} \right)^2 \right] dA - \rho \int_S \tilde{\phi}_i^e N_i^e \frac{\partial \tilde{\phi}}{\partial n} ds$$

La expresión final de I será la suma de los I en cada elemento

$$I(\tilde{\phi}) = \sum_1^N I^e(\tilde{\phi}^e)$$

4.- Finalmente, para encontrar los $\tilde{\phi}_i$ que minimizan la $I(\tilde{\phi})$, para cada elemento tomamos $\partial I^e / \partial \tilde{\phi}_i^e = 0$ obteniendo:

$$\tilde{\phi}_i^e \int_A \left[\frac{\partial N_i^e}{\partial x} \frac{\partial N_j^e}{\partial x} + \frac{\partial N_i^e}{\partial y} \frac{\partial N_j^e}{\partial y} \right] dA = \int_S N_j^e \frac{\partial \tilde{\phi}}{\partial n} ds$$

lo que hemos llamado $K_{ji}^e \tilde{\phi}_i^e = B_j^e$ para cada elemento y $K_{ji} \tilde{\phi}_i = B_j$ para el area total.

Coordenadas de Area

A pesar de que las integrales del tipo

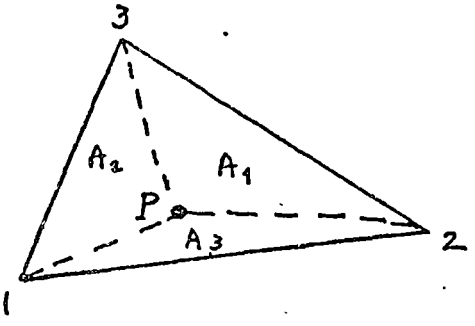
$$\int_A \frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x}\,dA$$

no encuentren ninguna dificultad conceptual para su integración ya que N_i es un polinomio en x y y de bajo orden, y dA es un elemento infinitesimal de área de un triángulo, el proceso de integración es largo y tedioso además de ser una excelente oportunidad para cometer errores algebraicos.

El uso de coordenadas de area, simplificará notablemente proceso y nos permitirá tabular en forma general varias integrales que aparecen con frecuencia en estos problemas.

Sea un triángulo de area A. Cualquier punto interior puede ser identificado inequívocamente indicando el area formada al unir P con los 3 vértices.

Por ejemplo:



el punto P en este triangulo queda identificado al indicar A_1, A_2, A_3. (Es importante mantener la notación aquí indicada, esto es A_1 frente al vértice 1, A_2 frente al vértice 2, A_3 frente al 3) Para hacer el sistema adimensional definimos las coordenadas de área como

$$\zeta_1 = \frac{A_1}{A} \qquad \zeta_2 = \frac{A_2}{A} \qquad \zeta_3 = \frac{A_3}{A} \qquad (A = \text{area del } \Delta_e)$$

Es claro que $\boxed{\zeta_1 + \zeta_2 + \zeta_3 = 1}$

Es inmediatamente aparente que las coordenadas de area del vértice o nudo 1. serán $\zeta = (1,0,0)$; para el punto 2 $\zeta = (0,1,0)$ y para 3 $\zeta = (0,0,1)$
Los lados del triangulo pueden ser facilmente descritos como:

lado	1-2 (l_3)	$\zeta_3 = 0$
lado	2-3 (l_1)	$\zeta_1 = 0$
lado	3-1 (l_2)	$\zeta_2 = 0$

Si queremos, por ejemplo, que ϕ varie linealmente [illegible] dentro del triángulo, anteriormente teníamos que al usar x y y quedaba como

$$\phi = \{ \phi_1 [(x_2 y_3 - x_3 y_2) + (y_2 - y_3) x + (x_3 - x_2) y] + \phi_2 [(x_3 y_1 - x_1 y_3) + (y_3 - y_1) x + (x_1 - x_3) y] + \phi_3 [(x_1 y_2 - x_2 y_1) + (y_1 - y_2) x + (x_2 - x_1) y] \} /$$

usando las coordenadas de area (ζ) la misma expresión se reduce a

$$\phi = \phi_1 \zeta_1 + \phi_2 \zeta_2 + \phi_3 \zeta_3$$

(que sencillo!!)

escrito en nuestra notación, en el elemento e

$$\phi^e = N_i^e \phi_i^e = N_1^e \phi_1^e + N_2^e \phi_2^e + N_3^e \phi_3^e$$
$$= \zeta_1^e \phi_1^e + \zeta_2^e \phi_2^e + \zeta_3^e \phi_3^e$$

Si queremos que una función varíe en forma cuadrática tendremos

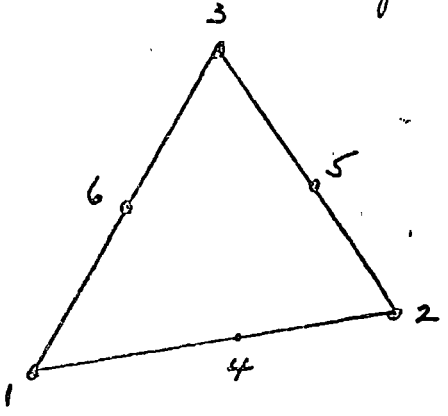
$$\phi = \phi_1 [\zeta_1 (2\zeta_1 - 1)] + \phi_2 [\zeta_2 (2\zeta_2 - 1)] + \phi_3 [\zeta_3 (2\zeta_3 - 1)] + \phi_4 \, 4\zeta_1 \zeta_2 + \phi_5 \, 4\zeta_2 \zeta_3 + \phi_6 \, 4\zeta_3 \zeta_1$$

lo que es relativamente fácil de probar.

(Nótese, y es importante, que al hacer una variación cuadrática aparecieron otras 3 nuevas constantes, ϕ_4 ϕ_5 y ϕ_6. Eso se debe a que una variación cuadrática en un area tiene 6 constantes, mientras que en una variación lineal solo tenía la mitad).

En la forma que definimos la variación cuadratica la numeración del triangulo será así:

(siempre contrareloj)



Sería interesante como ejercicio que verificara Ud que ϕ realmente toma el valor ϕ_i al reemplazar los ξ_i para ese punto i.

Es posible (y útil en este momento) establecer claramente cuál es la relación entre ξ y x y.

$$\begin{Bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{Bmatrix} = \frac{1}{2A} \begin{Bmatrix} x_2 y_3 - x_3 y_2 & y_2 - y_3 & x_3 - x_2 \\ x_3 y_1 - x_1 y_3 & y_3 - y_1 & x_1 - x_3 \\ x_1 y_2 - x_2 y_1 & y_1 - y_2 & x_2 - x_1 \end{Bmatrix} \begin{Bmatrix} 1 \\ x \\ y \end{Bmatrix}$$

(esto se cumple <u>siempre</u>, por definición de ξ, y <u>no</u> varía al hacer cuadrática o cúbica la variación).

Si definimos $a_k = x_j - x_i$ y

$b_k = y_i - y_j$, entonces

$2A = a_j b_i - a_i b_j$

De lo anterior se puede obtener:

$$\frac{\partial \zeta_i}{\partial x} = \frac{b_i}{2A} \quad \text{y} \quad \frac{\partial \zeta_i}{\partial y} = \frac{a_i}{2A}$$

Es fácil demostrar que

$$\int_A \zeta_i \, dA = \frac{A}{3} \; ; \; \int_A \zeta_i \zeta_j \, dA = \frac{A}{12} \; ; \; \int_A \zeta_i \zeta_i \, dA = \frac{A}{6}$$

$i \neq j$

(sería un buen ejercicio para el alumno verificar estas integrales).

Finite element simulation of water circulation in the North Sea

C. A. Brebbia and P. W. Partridge*

Department of Civil Engineering, University of Southampton, Southampton SO9 5NH, UK
(Received February 1976)

The modelling of tidal effects, storm surges and currents in large bodies of water is considered. The solution is attempted using the evolutionary shallow water equations with velocities and wave heights as unknowns. Two finite element simulation models are described based on six noded triangular elements. Special consideration has been given to the adequacy of the models which were applied to the North Sea only after extensive tests in channels. Results for velocities and wave heights are compared and discussed. A set of conclusions on the applicability and scope of the models is presented.

Introduction

This paper is concerned with the modelling of tidal effects, storm surges and current patterns in large bodies of water. The solution is attempted using the shallow water equations, which are evolutionary equations with velocities and wave heights as unknowns. They require the initial conditions as well as the boundary conditions to be known.

The solution of these equations is usually found by applying a numerical technique. The method used is of fundamental importance. In a finite element or finite difference approach the grid size will determine the type of phenomenon which can be investigated. In addition grid size relates to stability criterion and accuracy in evolutionary problems.

The refinement of a model, though desirable in principle, may demand a large number of parameters which require more experimental data. These data can be difficult to obtain and produce a new type of error affecting the confidence one can have in the results. The analyst usually has to compromise between having a sophisticated model or a practical one, giving reliable results for the variables under consideration. In addition, large models are expensive to run.

We describe here two finite element models. Both models have been developed using six noded triangular elements, but one is based on an implicit integration scheme, the other in an explicit one (the former allows for elements with curved sides). Special consideration has been given to the adequacy of the models and only after extensive tests on channels[1] were they applied to North Sea studies. The North Sea is an important and busy seaway, especially since the discovery of gas and oil. From the numerical point of view, the area is well-conditioned, North Sea topography being regular and changes in depth gradual. Nevertheless, the models can and have been applied to different regions (e.g., the Solent in England[2]. Other models of the North Sea exist: an explicit finite difference scheme by Heaps[3], an implicit three-node finite element one by Grotkop[4] and a quartic quadrilateral finite element model by Davis and Taylor[5].

The present model is based on the shallow water equations which are vertically averaged versions of Navier–Stokes equations, and take into consideration tides, bottom friction, advective forces, coriolis, wind tangential stresses and atmospheric pressure gradients.

Results for velocities and wave heights in the North Sea are compared and discussed. A set of conclusions on the applicability and scope of the models is presented, indicating areas where further work is required.

Shallow water equations

The evolutionary equations used in marine and certain types of estuarial modelling are called the shallow water equations. They are a vertically integrated version of Navier–Stokes momentum equations and the continuity equation which acts as a constraint condition. In addition initial and boundary conditions have to be fulfilled. The different assumptions involved

*Present address: Federal University, Porto Alegre, Brazil

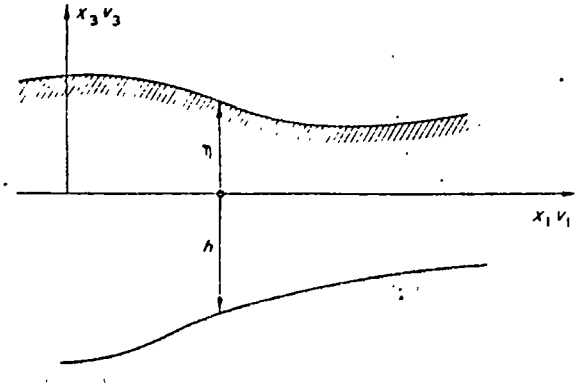


Figure 1 Geometrical notation for the shallow water equations

are treated in detail elsewhere[2,6]. The two shallow water momentum equations are:

$$\frac{\partial V_1}{\partial t} + V_1\frac{\partial V_1}{\partial x_1} + V_2\frac{\partial V_1}{\partial x_2} = B_1$$
$$\frac{\partial V_2}{\partial t} + V_1\frac{\partial V_2}{\partial x_1} + V_2\frac{\partial V_2}{\partial x_2} = B_2 \qquad (1)$$

where

$$B_1 = \Omega V_2 - g\frac{\partial \eta}{\partial x_1} - \frac{\partial}{\partial x_1}\left(\frac{p_a}{\rho}\right) - \frac{1}{\rho}\tau_1\Big|_b + \frac{1}{\rho}\tau_1\Big|_s \qquad (2)$$
$$B_2 = -\Omega V_1 - g\frac{\partial \eta}{\partial x_2} - \frac{\partial}{\partial x_2}\left(\frac{p_a}{\rho}\right) - \frac{1}{\rho}\tau_2\Big|_b + \frac{1}{\rho}\tau_2\Big|_s$$

V_i are the averaged velocities:

$$V_i = \frac{1}{H}\int_{-h}^{\eta} v_i \, dx_3 \qquad (3)$$

H is the total depth, $H = \eta + h$, where η is the wave height above a certain datum plane and h is the depth from the datum to the bottom of the sea. x_3 is the coordinate in the vertical direction (*Figure 1*), $\Omega = 2\omega \sin\phi$ is the coriolis coefficient, ϕ is the latitude and ω the angular rotation of the earth, g is gravity, ρ the water density and p_a the atmospheric pressure. The surface and bottom stresses are written as:

$$\tau_i\Big|_s = \frac{\gamma}{\rho}\frac{W_i}{H^2}(W_1^2 + W_2^2)^{1/2} \qquad i = 1;2$$
$$\tau_i\Big|_b = -\left(\frac{g}{c^2}\right)\rho\frac{V_i}{H}(V_1^2 + V_2^2)^{1/2} \qquad i = 1,2 \qquad (4)$$

c is the Chezy coefficient, W_i are the wind speed components and γ is a parameter related to atmospheric density ρ_a (usually given as a constant multiplied by ρ_a).

In addition equations (1) have to satisfy the vertically integrated continuity equation, i.e.,

$$\frac{\partial H}{\partial t} + \frac{\partial}{\partial x_1}(HV_1) + \frac{\partial}{\partial x_2}(HV_2) = 0 \qquad (5)$$

The systems of equations (1) and (5) describe the movement of large bodies of shallow water. The factors affecting the movement are many: the morphology and position of the seabed, the shape and variations in shape of the coastline, friction between the seabed and the water, hence the material of the seabed, the meteorological conditions, including wind, etc. Although the circulation of the earth and the astronomical forces of the sun and the moon act on the water as body forces, the main cause of tidal water movements in areas such as the North Sea is the driving force caused by tidal motion of the water on the boundaries of the area under consideration.

The shape of the land surface containing the body of water is usually very complex, in some cases not even static, though the effects of erosion generally occur over too large a period of time to be important.

Bottom friction is introduced in the model via Chezy coefficients. The inadequacy of using constant Chezy coefficients for all the model is evident. The different materials making up the seabed have different frictional resistances as the water depth and the velocities change. It must be pointed out that bottom friction and wind are of great importance in the movement of shallow water.

The main causes of inaccuracies in tidal predictions are the wind forces and atmospheric pressure variations, which are important for large areas such as the North Sea.

Boundary and initial conditions

The solution of equations (1) and (5) require the knowledge of the corresponding boundary and initial conditions. The boundary conditions of the model are of two types: (a) fixed or land boundaries such as those given by the coastlines, where the normal velocities are zero, and the tangent velocity can be set free; (b) open boundaries where the elevation of the sea level (or the normal component of velocity) is prescribed.

The determination of the initial conditions requires the knowledge of the free surface position at $t = 0$. Usually this knowledge is not possible and the models have to be started with zero elevation and zero velocity conditions. This is called a 'cold start'.

Finite element model

In order to build finite element models the two momentum equations (1) and continuity (5), including influx-type boundary conditions have to be written in the following weighted residual way:

$$\iint\left\{\frac{\partial V_1}{\partial t} + V_1\frac{\partial V_1}{\partial x_1} + V_2\frac{\partial V_1}{\partial x_2} - B_1\right\}\delta V_1 \, dA = 0$$
$$\iint\left\{\frac{\partial V_2}{\partial t} + V_1\frac{\partial V_2}{\partial x_1} + V_2\frac{\partial V_2}{\partial x_2} - B_2\right\}\delta V_2 \, dA = 0 \qquad (6)$$
$$\iint\left\{\frac{\partial H}{\partial t} + \frac{\partial}{\partial x_1}(HV_1) + \frac{\partial}{\partial x_2}(HV_2)\right\}\delta H \, dA = \int(HV_n - H\bar{V}_n)\delta H \, dS$$

The continuity equation is usually integrated by parts to render a simpler expression. This integration gives:

$$\iint\left\{HV_1\frac{\partial \delta H}{\partial x_1} + HV_2\frac{\partial \delta H}{\partial x_2} - \frac{\partial H}{\partial t}\delta H\right\} dA = \int H\bar{V}_n\delta H \, dS \qquad (7)$$

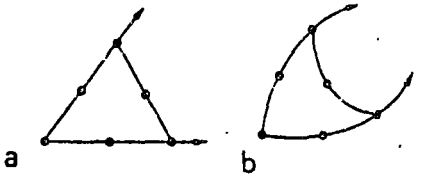


Figure 2 Six node elements (a) Straight sides, (b) curved sides

The above weighted residual statements [equations (6) and (7)] are the starting point for the finite element models. Assume that over an element the same interpolation function applies for the V_1, V_2 and H unknowns, i.e.

$$V_1 = \phi V_1^n, \qquad V_2 = \phi V_2^n, \qquad H = \phi H^n \tag{8}$$

ϕ is the interpolation function and V_1^n, H^n are nodal values of V_i, H.

In what follows six noded triangular finite elements with curved boundaries were used in order to define the boundaries better (*Figure 2*). These elements are called isoparametric and can be formulated by a simple coordinate transformation, the details of which have been given by Connor and Brebbia[7]. Curved elements have the important feature that they tend to eliminate the spurious forces that may be generated on the boundaries by straight side elements joining at an angle.

Substituting equations (8) into (6) and (7), one obtains:

$$\begin{aligned} M\dot{V}_1^n + KV_1^n - \Omega MV_2^n + G_1H^n + F_1 = 0 \\ M\dot{V}_2^n + \Omega MV_1^n + KV_2^n + G_2H^n + F_2 = 0 \end{aligned} \tag{9}$$

and

$$M\dot{H}^n - C_1V_1^n - C_2V_2^n + F_H = 0$$

where

$$K = \int \phi^T\phi_{,1}V_1\,\mathrm{d}A + \int \phi^T\phi_{,2}V_2\,\mathrm{d}A + \left(\frac{g}{c^2}\right)\int \phi^T\frac{(V_1^2+V_2^2)^{1/2}}{H}\phi\,\mathrm{d}A$$

$$G_i = g\int \phi^T\phi_{,i}\,\mathrm{d}A \qquad M = \int \phi^T\phi\,\mathrm{d}A$$

$$F_i = \int \phi^T\left(\frac{p_a}{\rho}\right)_{,i}\mathrm{d}A + \left(\frac{\gamma}{\rho}\right)\int \phi^T\frac{W_i}{H}(W_1^2+W_2^2)^{1/2}\,\mathrm{d}A \qquad i = 1,2$$

$$C_i = \int \phi_{,i}^T H\phi\,\mathrm{d}A, \qquad F_H = \int H\bar{V}_n\phi^T\,\mathrm{d}A$$

and

$$(\,)_{,i} = \frac{\partial}{\partial x_i}, \qquad \dot{(\,)} = \frac{\partial}{\partial t}$$

Equations (9) can be written as

$$\begin{bmatrix} M & \cdot & \cdot \\ \cdot & M & \cdot \\ \cdot & \cdot & M \end{bmatrix}\begin{Bmatrix} \dot{V}_1^n \\ \dot{V}_2^n \\ \dot{H}^n \end{Bmatrix} + \begin{bmatrix} K & -\Omega M & G_1 \\ \Omega M & K & G_2 \\ -C_1 & -C_2 & 0 \end{bmatrix}\begin{Bmatrix} V_1^n \\ V_2^n \\ H^n \end{Bmatrix} +$$

$$\begin{Bmatrix} F_1 \\ F_2 \\ F_H \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix} \tag{10}$$

or more simply,

$$M\dot{Q} + KQ = F \tag{11}$$

Formula (11) is valid for each unconnected element. The next stage is to assemble all the element equations into a global system and impose boundary conditions in H and V_i. To eliminate proliferation of notation the global system will be defined with the same notation as equation (11).

Time integration

Two time integration schemes were used, one an implicit and the other an explicit scheme. The implicit integration procedure is the trapezoidal rule. Starting with:

$$M\dot{Q} + KQ = F \tag{12}$$

one assumes:

$$\dot{Q} = \frac{Q_t - Q_0}{\Delta t} \qquad Q = \frac{Q_t + Q_0}{2}$$
$$F = \frac{F_0 + F_t}{2} \tag{13}$$

Hence equation (12) becomes:

$$\left(\frac{2}{\Delta t}M + K\right)Q_t = (F_0 + F_t) + \left(\frac{2}{\Delta t}M - K\right)Q_0 \tag{14}$$

or

$$K^*Q_t = F^* \tag{15}$$

The recurrence relationship is then:

$$Q_t = (K^*)^{-1}F^* \tag{16}$$

The K^* matrix to be inverted generally is a large non-symmetric banded matrix of size approximately three times the number of nodes by six times the element band width (i.e., the maximum difference between element nodal point numbers plus one). The computer program has been optimized by taking boundary conditions into account in such a way that the corresponding rows and columns are eliminated from the element matrices before assembling. This significantly reduces the maximum size of the global matrix. It was also advantageous to store the matrix in a one-dimensional form such that only one and not two addresses need to be evaluated each time an element of the array is accessed.

The explicit time integration scheme used was the well-known fourth order Runge–Kutta scheme.

North Sea model

The above finite element formulation has been applied to model the North Sea. This is a shallow sea varying in depth from under 50 m in the south to 400 m in a trench off the coast of Norway. Depths were obtained from Admiralty charts. Sections were drawn at different angles across the whole region to determine

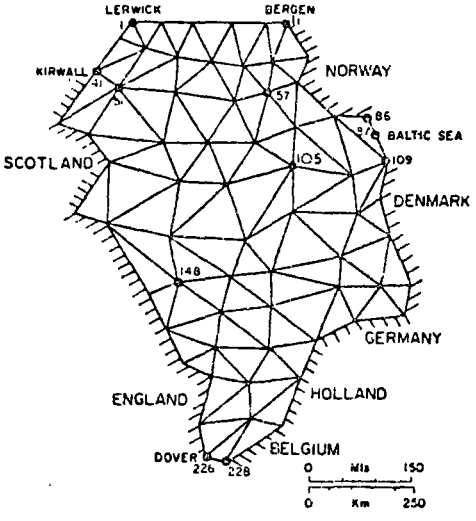


Figure 3 Finite element mesh for the North Sea

the best locations for nodes and the more accurate way of representing the bottom topography. Elements were carefully positioned in order to obtain the best possible representation of the topography using a predetermined number of nodes. The final mesh comprises 228 nodes and 97 six noded elements as shown in *Figure 3*.

The tidal characteristics of the North Sea are complicated, tidal amplitudes vary from zero to six metres and high water times change throughout the cycle around each of the three amphidromic points.

Tidal heights for boundary conditions were taken from the charts of co-tidal lines. The waveheight forcing functions were specified in the form (*Figure 4*)

$$\bar{\eta} = a_0\left[\sin\left(\frac{2\pi t}{T} + \varepsilon\right) + 1\right] \tag{17}$$

on each of the extremes of the tidal boundaries, with the intermediate heights being linearly interpolated. This formulation approximates the most important tidal component for the North Sea. These curves have been taken from the Admiralty Tide Tables They are assumed to be referred to the same datum since other information is not available. Because of this, the results presented in this paper may not be quantitatively correct but the comparison between models are still valid. The charts of co-tidal lines give the approximate tidal range and high water times for the interior points on the grid. The circulation pattern and velocity magnitudes for parts of the Sea may be seen in the tidal stream atlases[8,9].

Wind and storm surges were not modelled as this series of tests was carried out to investigate the general performance of the model.

After several tests it was decided to take a continuous tidal boundary in the Northern part from node 41 to 1 and 1 to 11 (*Figure 4*), otherwise instability originated from the Shetland Islands element, showing that one element is inadequate to represent a discontinuous tidal boundary properly. Tidal conditions were also specified at the Dover Strait (points 226 to 228) and after a number of trials also for the Baltic Sea boundary (nodes 86–87–109). It was found that inclusion of the Baltic Sea improved the waveheight results.

On the land boundaries, the no-slip boundary condition $V_1 = V_2 = 0$ is specified. This assumption simplifies the necessary computing and is reasonable as the North Sea is a regularly shaped region. Because of the imposition of this condition, curved boundary elements, which are more expensive to run, were not necessary. By contrast when modelling the Solent, curved sided boundary elements were used allowing the tangential velocity to remain free[2].

Stability and accuracy

The smallest stability limit for the North Sea as given by the Friedricks–Lewy–Courant condition, occurs for an element off the Norwegian coast. It gives:

$$\Delta t \leqslant 450\,\text{sec} \tag{18}$$

The worst case on a tidal boundary gives $\Delta t \leqslant 650$ sec. This is important as instabilities always start at these boundaries. The average value is around 900 sec and for the shallow southern North Sea the criterion suggests a limiting time step of less than 2000 sec. The explicit programme which uses a fourth-order Runge–Kutta procedure was run with a time step of 600sec. For the implicit programme instead a time step of 30min was used.

To obtain stable results with both models requires the application of special techniques. For this work three different techniques were used. The first and simplest of them is to work always with a constant value of friction over all the region, starting with a low value ($c = 10\,\text{m}^{1/2}/\text{sec}$) and increasing it by 10 over two to four cycles. This technique did not give good results and the solution tends to become unstable for large values of c, i.e., small values of friction. In addition it is unrealistic to assume that the Chezy coefficient will be the same over all the domain.

The second technique was to prescribe a higher order of friction for the elements on tidal boundaries and a smaller value for internal elements. This is because there is a general tendency for the tidal boundary to generate disturbances. These perturbations may be due to a number of factors and cause the transmission of short waves through the system. The specification of higher values of friction for the tidal boundary elements reduce the propagation of errors.

It was decided to apply a Chezy coefficient $c = 20\,\text{m}^{1/2}/\text{sec}$ at the boundaries, which produces stable results.

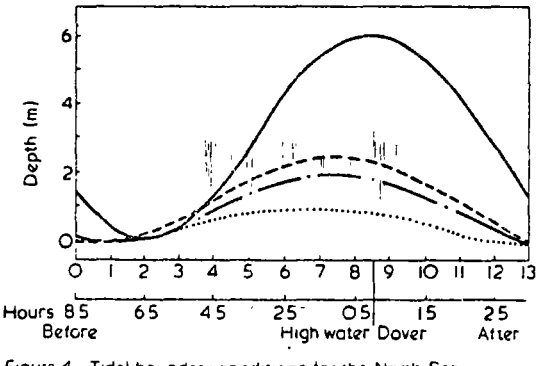


Figure 4 Tidal boundary conditions for the North Sea
Bergen, — — —, Lerwick, – – – –, Kirkwall ———, Dover

The third different stabilizing technique is to start with a realistic value of friction from the beginning and try to remove the short waves by 'numerical smoothing'. This smoothing can also be applied every time the level of friction is decreased. The operation needs to be carried out for a number of steps and after the perturbations have been removed the solution does not need to be smoothed any longer. This numerical smoothing consists of taking for the next time step not the actual nodal values just obtained but weighted averages. These averages are calculated by weighting every nodal value by a constant and adding to it the weighted values of the neighbouring points. For the programmes described here the node under consideration has half the weight and the other half is distributed among six neighbouring points proportionally to their area of influence. (For boundary nodes only 3 neighbouring points are taken into consideration.) In general, the coarser the mesh the more relative weight the central node will have.

The actual weight used does not seem to be too important provided that the coefficient for the node under consideration is reasonably large by comparison with the coefficients for the neighbouring nodes A simpler way of weighting may be for instance, to multiply the solution vector by the mass matrix. Smoothing has been successfully applied by the authors in small estuarial areas, such as the Solent in England but it is less necessary for the North Sea as the system is more stable.

Tests

Many tests were run with different time integration schemes, friction coefficients and smoothing schemes but only two of them will be presented for brevity. The first, Test 1, uses implicit integration ($\Delta t = 30$ min) and the second, Test 2, explicit fourth-order Runge–Kutta ($\Delta t = 10$ min).

Test 1 was started with a Chezy coefficient of $10\,m^{1/2}$/sec, after two tidal cycles this was increased to 15 and after another two to 20. Then the friction 20 was left for the elements on the tidal boundary but the value of the internal friction was decreased to $c = 40$ over four tidal cycles. Finally, the internal Chezy coefficient was taken as a variable given by:

$$c = 15 \log_e(0.9H) \qquad [\text{in } m^{1/2}/\text{sec}] \qquad (19)$$

This formula gives low friction in the interior of the North Sea (e.g. for $H = 55$ m, $c = 60$). The same friction was applied during 6 more tidal cycles to obtain repetition of results. In addition the results were numerically smoothed over 3 h (i.e., 6 steps) after change of Chezy's coefficient, then the smoothing was stopped. For the level of friction given by formula (19) the velocity ellipses tend to increase in size and their drift is accentuated as the level of friction is reduced. As the friction is variable, the results are not being obtained under constant conditions of damping and the ellipses do not quite close even after three or four tidal cycles. It is surprising that good results were reported by Davis and Taylor[7] after only three tidal cycles from cold start, using this variable friction formula.

Test 2 was run with tidal boundary elements friction

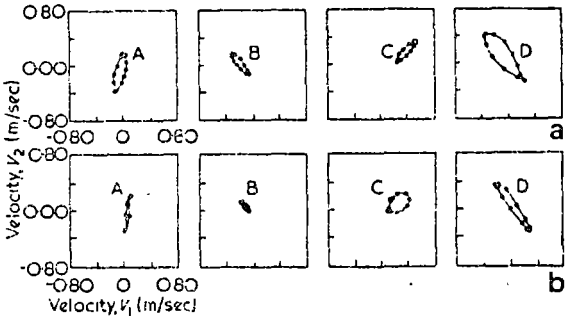


Figure 5 Velocity ellipses for (a) implicit and (b) explicit models A, Node 51, B node 57, C, node 105, D, node 148

$c = 20$ ($m^{1/2}$/sec) and values of internal friction of 40 and 60. The solution was initiated from the implicit model results for $c = 20$ throughout instead of direct cold start. (This was done simply to save computer costs.)

Results

Velocity ellipses are useful to find out if steady state has been reached, if there are any disturbances present in the system which are likely to cause instability, the magnitude of drift velocities, the changes in velocity magnitudes due to changes in Chezy's coefficients, time step, etc.

Two comparable sets of results obtained using the same tidal boundary Chezy coefficient ($20\,m^{1/2}$/sec) and the same internal friction ($c = 40$) are shown in *Figure 5*. It can be seen that the ellipses for the explicit scheme tend to be smaller (this is also true for other points) than those obtained with implicit integration The drift is also less in the explicit method. This may be due to the large time step used in the implicit solution. As a smoother circulation pattern and smaller drifts are obtained one could conclude that the explicit solution is more accurate in this case. There are also slight differences in the shape of the ellipses.

The large drift (in the order of 10 to 20 cm/sec for some points) may also be explained by the coarseness of the grid. A typical plot of velocities over all the Sea is also shown in *Figure 6* (results are from the implicit programme, the Chezy coefficient is 20 on tidal boundaries and 40 inside), where they are compared against results published in the tidal stream atlases available in Britain. The general trend of the velocities compares well.

We should be aware, however, that the tidal stream atlases velocities are smoothed out, the observations are made in only the top layer of the water. The programme instead yields depth averaged velocities given by equation (3). Hence the currents worked out by the programme are not exactly as the atlases currents The programme results are being affected by local fluctuations in depth to a greater degree than the figures in the atlases

Graphs showing the waveheight solutions for the same tests and at the same points as the velocities are shown in *Figure 7*. There are comparatively high variations in tidal range and high water times. The flatness of the waveheight curves at nodes near an amphidromic point is also noticeable.

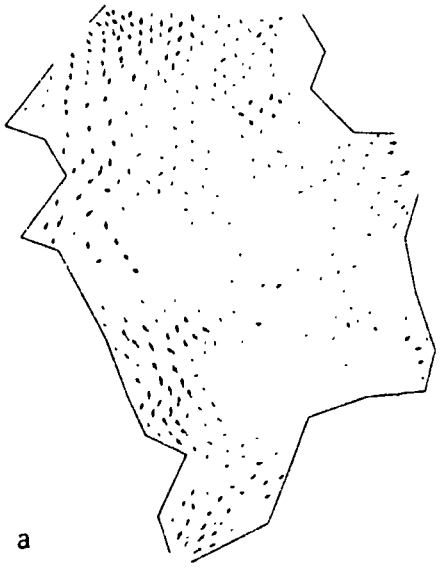


Figure 6a Velocity vectors obtained with the implicit programme 2 hours after high water at Dover

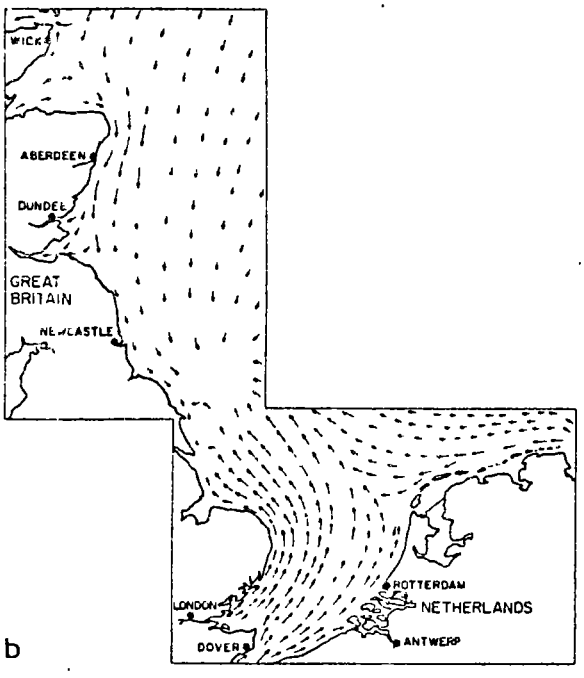


Figure 6b Velocity vectors from the Admiralty charts 2 hours after high water

Finally the co-range lines chart for the North Sea has been computed from results obtained using the implicit programme with friction $c = 20\,m^{1/2}$ sec on tidal boundary and for interior elements c as given by formula (19) and shown in *Figure 8*. If the co-range lines are compared against the results shown in Admiralty chart 5058 the agreement is reasonable. Similar curves were also reported by Grotkop[4] and Nihoul[6].

Conclusions

For a constant level of friction throughout the grid the results became unstable for a value of Chezy's coefficient of $60\,m^{1/2}$/sec unless special procedures are used. Sudden reductions of the level of friction (i.e., increases in c) cause a small shock to be transmitted through the system. Short waves are also generated on the tidal boundary by the discrete changes in the imposed tidal height at each time step. A way of damping out those effects is by decreasing the c coefficients in elements on the tidal boundary and this technique has been used for all the results shown in this paper. Another damping procedure is the numerical smoothing, for which the model could be started with a realistic level of friction and the results numerically smoothed until the disturbance caused by

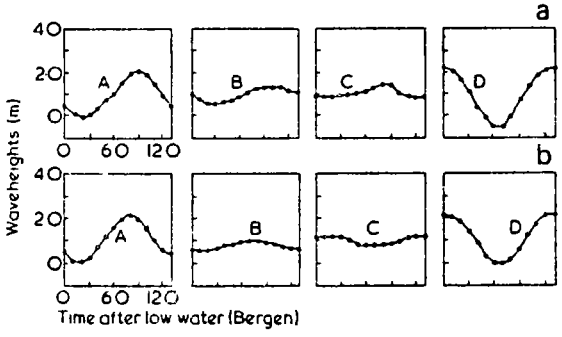


Figure 7 Wavelength for (a) implicit and (b) explicit models A. Point 51, B. point 57, C. point 105, D. point 148

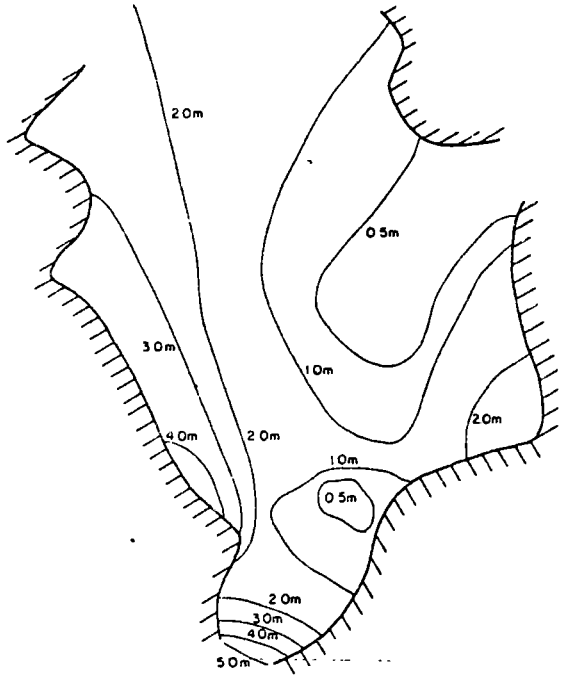


Figure 8 Co-range tidal lines

the cold start has been reduced. Since it is not necessary to alter the friction parameter after this the introduction of further short waves may in some cases be avoided. For the results reported in this paper the numerical smoothing technique was only used once for a short time, to stabilize the implicit model results when the Chezy coefficient changed from 40 to the value given by formula (19).

Both time integration schemes, implicit and explicit, give similar results, however, the fourth-order Runge–Kutta is more accurate using $\Delta t \doteq 10$ min which is the highest allowable time step in this case. But it should be also pointed out that the implicit programme ($\Delta t = 30$ min) needs less than half the computer time required for the explicit programme.

For evolutionary processes of the type here described computer time can be very expensive and the programme should be further optimized before undertaking production runs. It seems, however, that for a problem with the dimensions of the North Sea and a finite element grid similar to the one used here, the explicit programme may be more convenient to use. This is not the case when the domain is smaller (for instance for the Solent) or the grid very fine. In other words implicit schemes are more expensive per timestep than explicit ones but allow for larger timesteps. This can be of interest in problems where the time step may be increased considerably over a simpler scheme, without significantly affecting the accuracy of the results

The viability of the six nodes finite element circulation model for the North Sea has been established. Second order elements of this type are especially suitable to represent accurately the topography of the region (i.e., variable depth and curved boundaries).

References

1 Partridge, P. W. and Brebbia, C. A. *Proc. Am. Soc. Mech. Eng (J. Hydraulics Div.)* 1976, in press
2 Brebbia, C. A. and Partridge, P. W., in 'Mathematical Models for Environmental Problems', (Ed. C. A. Brebbia) Pentech Press, London, 1976, p 141
3 Heaps, N. S. *Mem. Soc. R. Sci., Liege*, 1972, 6, 143
4 Grotkop, G. *Comput. Methods Appl. Mech. Eng.*, 1973, 2
5 Davis, J. M. and Taylor, C. *Proc. Finite Element Methods in Flow Problems Conf., Swansea*, 1973
6 Nihoul, J. C. J. (Ed.), 'Modelling of Marine Systems', Elsevier, Amsterdam, 1975
7 Connor, J. J. and Brebbia, C. A., 'Finite Elements for Fluid Flow', Butterworths, London, 1976
8 Hydrographer for the Navy, 'Tidal Streams Atlas North Sea — Southern Portion', HMSO, London, 1962
9 Hydrographer of the Navy, 'Tidal Streams Atlas North Sea — Flamborough Head to Portland Forth', HMSO, London, 1962

USOS DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPRESION EN AGUAS, COSTERAS LAGOS Y RIOS

CIRCULACION EN LAGOS

MAYO, 1978.

JOURNAL OF THE
HYDRAULICS DIVISION

TRANSIENT FINITE ELEMENT SHALLOW LAKE CIRCULATION

By Frank D. L. Young[1] and James A. Liggett,[2] M. ASCE

INTRODUCTION

The general problem of lake circulation is very complex and remains unsolved for practical engineering needs; wide use, however, has been made in recent years of the steady-state solution to the circulation of a nonstratified lake. A finite element program described in Ref. 3 has had considerable utilization in many parts of the world. This paper extends that analysis to transient problems.

A steady-state analysis uses the Ekman generalizations of Welander (11) under the restrictions of constant vertical eddy viscosity, no horizontal eddy viscosity, and a small Rossby number. Under these conditions the three-dimensional velocity field can be found in an efficient manner since the depthwise variation of the horizontal velocities is removed through integration in the vertical. A finite difference program using this technique is described in Ref. 6. The finite element program (3) was written to take advantage of the finite element network's ability to represent odd geometries and for the ease of universal use, including simplified input data and mesh generation. Considerable success has been reported.

A transient solution, described in Ref. 4, exists. There the vertical velocity variation was removed by a Fourier transform, but the evaluation of a considerable number of Fourier series terms each time step proved to be inefficient. A full three-dimensional velocity analysis (5) was as efficient and more general. The method described herein uses a Laplace transform with numerical inversion. The technique of removing the vertical velocity distribution fits easily into the method; also, the program does not step through time in the finite difference sense but needs to be run only about six to 12 times in a typical analysis. The gain in ease in handling such problems is very large, as is the increase in efficiency.

Techniques frequently used for time-dependent problems within the framework

Note.—Discussion open until July 1, 1977. To extend the closing date one month, a written request must be filed with the Editor of Technical Publications, ASCE. This paper is part of the copyrighted Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, Vol. 103, No. HY2, February, 1977. Manuscript was submitted for review for possible publication on April 14, 1976.

[1]Research Assoc., School of Civ. and Environmental Engrg., Cornell Univ., Ithaca, N.Y.

[2]Prof., School of Civ. and Environmental Engrg., Cornell Univ., Ithaca, N.Y.

of the finite element method include Runge-Kutta integration (1), finite differences in time (10), and finite elements in time (2). An interesting example of the step method used to solve a practical problem is cited in Ref. 7. In that calculation, time steps of 2 min were used to simulate up to 4 hr of real time. In addition, an iterative method was used. The number of solutions using the Laplace transform technique is an order of magnitude less. Unfortunately, algebraic development of the present scheme is difficult and lengthy, even though the problem is conceptually straightforward. In order to avoid stumbling over the algebra, many of the expressions have been placed in Appendix I.

Governing Equations

Basic assumptions and derivation of the governing equations are exactly the same as in Ref. 4 and will not be repeated herein. The equations are

$$\frac{\partial u}{\partial t} - fv = -\frac{1}{\rho}\frac{\partial p}{\partial x} + \eta\frac{\partial^2 u}{\partial z^2} \quad \ldots (1)$$

$$\frac{\partial v}{\partial t} + fu = -\frac{1}{\rho}\frac{\partial p}{\partial y} + \eta\frac{\partial^2 v}{\partial z^2} \quad \ldots (2)$$

$$g = -\frac{1}{\rho}\frac{\partial p}{\partial z} \quad \ldots (3)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad \ldots (4)$$

Subjected to the boundary conditions

$$u = v = w = 0 \quad \text{on all solid boundaries.} \quad z = -h \quad \ldots (5)$$

$$\text{and} \quad \eta\frac{\partial u}{\partial z} = \tau_x; \quad \eta\frac{\partial v}{\partial z} = \tau_y \quad \text{at the free surface,} \quad z = 0 \quad \ldots (6)$$

The notation is defined in Ref. 4 and is included in Appendix III. The following nondimensional parameters are defined:

$$x^* = \frac{x}{L}; \quad y^* = \frac{y}{L}; \quad z^* = \frac{z}{D}; \quad t^* = ft; \quad u^* = \left(\frac{fL}{gD}\right)u;$$

$$v^* = \left(\frac{fL}{gD}\right)v; \quad w^* = \left(\frac{fL^2}{gD^2}\right)w; \quad h^* = \frac{h}{D}; \quad p^* = \left(\frac{p}{\rho gD}\right) + \left(\frac{z}{D}\right);$$

$$\Delta^* = \left[\frac{fL\tau_x}{\eta g}\right]; \quad \text{and} \quad \Gamma^* = \left[\frac{fL\tau_y}{\eta g}\right].$$

These variables are introduced into the equations; and the asterisks are omitted Hereafter, only dimensionless variables appear (except in the definition of the Taylor and Ekman numbers). The dimensionless equations are

$$\frac{\partial u}{\partial t} - v = -\frac{\partial p}{\partial x} + \frac{1}{2m^2}\frac{\partial^2 u}{\partial z^2} \quad \ldots (7)$$

$$\frac{\partial v}{\partial t} + u = -\frac{\partial p}{\partial y} + \frac{1}{2m^2}\frac{\partial^2 v}{\partial z^2} \quad \ldots (8)$$

$$\frac{\partial p}{\partial z} = 0 \quad \ldots (9)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad \ldots (10)$$

With the boundary conditions

$$u = v = w = 0 \quad \text{at} \quad z = -h \quad \ldots (11)$$

$$\text{and} \quad \frac{\partial u}{\partial z} = \Delta; \quad \frac{\partial v}{\partial z} = \Gamma \quad \text{at} \quad z = 0 \quad \ldots (12)$$

To this point, the details can be found in Ref. 4. The problem is described in terms of three parameters: m (in which $2m^2 = fD^2/\eta = Ta$, the Taylor number, or $1/2m^2 = E$, the Ekman number); Δ; and Γ. In addition, depth is a prescribed function of the horizontal coordinates, $h = h(x, y)$.

Solution

The time derivatives of the equations are removed by means of the Laplace transform. The transformed variables are

$$\bar{u} = \int_0^\infty ue^{-st}\,dt \quad \ldots (13)$$

Thus all terms are multiplied by e^{-st} and integrated to yield

$$s\bar{u} - \bar{v} = -\frac{\partial \bar{p}}{\partial x} + \frac{1}{2m^2}\frac{\partial^2 \bar{u}}{\partial z^2} \quad \ldots (14)$$

$$s\bar{v} + \bar{u} = -\frac{\partial \bar{p}}{\partial y} + \frac{1}{2m^2}\frac{\partial^2 \bar{v}}{\partial z^2} \quad \ldots (15)$$

$$\frac{\partial \bar{p}}{\partial z} = 0 \quad \ldots (16)$$

$$\frac{\partial \bar{u}}{\partial x} + \frac{\partial \bar{v}}{\partial y} + \frac{\partial \bar{w}}{\partial z} = 0 \quad \ldots (17)$$

with boundary conditions

$$\bar{u} = \bar{v} = \bar{w} = 0 \quad \text{at} \quad z = -h \quad \ldots (18)$$

$$\text{and} \quad \frac{\partial \bar{u}}{\partial z} = \bar{\Delta} \quad \frac{\partial \bar{v}}{\partial z} = \bar{\Gamma} \quad \text{at} \quad z = 0 \quad \ldots (19)$$

Also, initial conditions of $u = v = 0$ have been used in Eqs. 14 and 15. If other initial conditions are desired, additional terms would be included in Eqs. 14 and 15 which are $u(x, y, z, 0)$ and $v(x, y, z, 0)$.

Since Eq. 16 states that $\bar{p}$ is not a function of z, Eqs. 14 and 15 can be vertically integrated, giving

$$\bar{u} = \frac{-1}{s^2+1}\left(s\frac{\partial \bar{p}}{\partial x} + \frac{\partial \bar{p}}{\partial y}\right) + \cos Nz(c_2 e^{Mz} - c_4 e^{-Mz})$$

$$- \sin Nz(c_1 e^{Mz} - c_3 e^{-Mz}) \quad \ldots \quad (20)$$

$$\text{and} \quad \bar{v} = \frac{-1}{s^2+1}\left(-\frac{\partial \bar{p}}{\partial x} + s\frac{\partial \bar{p}}{\partial y}\right) + \cos Nz(c_1 e^{Mz} + c_3 e^{-Mz})$$

$$+ \sin Nz(c_2 e^{Mz} + c_4 e^{-Mz}) \quad \ldots \quad (21)$$

$$\text{in which} \quad M = \sqrt{2}\, m \sqrt{R} \cos\frac{\phi}{2} \quad \ldots \quad (22)$$

$$N = \sqrt{2}\, m \sqrt{R} \sin\frac{\phi}{2} \quad \ldots \quad (23)$$

$$R = \sqrt{(1+s^2)} \quad \ldots \quad (24)$$

$$\phi = \tan^{-1}\frac{1}{s} \quad \ldots \quad (25)$$

The constants, c_1, c_2, c_3, c_4, can be determined through the boundary conditions and are found to be

$$c_1 = c_3 + \frac{1}{(M^2+N^2)}(M\bar{\Gamma} - N\bar{\Delta}) \quad \ldots \quad (26)$$

$$c_2 = -c_4 + \frac{1}{(M^2+N^2)}(N\bar{\Gamma} + M\bar{\Delta}) \quad \ldots \quad (27)$$

$$c_3 = \frac{1}{\alpha}\left[\frac{\gamma s - \epsilon}{s^2+1}\frac{\partial \bar{p}}{\partial x} + \frac{\gamma + \epsilon s}{s^2+1}\frac{\partial \bar{p}}{\partial y} + (\beta\epsilon - \delta\gamma)\bar{\Gamma} + (\beta\gamma + \delta\epsilon)\bar{\Delta}\right] \quad \ldots \quad (28)$$

$$c_4 = \frac{1}{\alpha}\left[\frac{-(\gamma + \epsilon s)}{s^2+1}\frac{\partial \bar{p}}{\partial x} + \frac{\gamma s - \epsilon}{s^2+1}\frac{\partial \bar{p}}{\partial y} + (\gamma\beta + \epsilon\delta)\bar{\Gamma} + (\gamma\delta - \epsilon\beta)\bar{\Delta}\right] \quad \ldots \quad (29)$$

Note for later reference (Eq. 55) that $\partial\bar{p}/\partial x$, $\partial\bar{p}/\partial y$, c_1, ..., c_4 all approach infinity as $1/s$ as $s \to 0$. The abbreviations previously used (and in Appendix I) are

$$\alpha = \cos^2 Nh(e^{-Mh} + e^{Mh})^2 + \sin^2 Nh(e^{-Mh} - e^{Mh})^2 \quad \ldots \quad (30)$$

$$\beta = \frac{e^{-Mh}}{(M^2+N^2)}(N \sin Nh - M \cos Nh) \quad \ldots \quad (31)$$

$$\gamma = \sin Nh(e^{-Mh} - e^{Mh}) \quad \ldots \quad (32)$$

$$\delta = \frac{e^{-Mh}}{(M^2+N^2)}(M \sin Nh + N \cos Nh) \quad \ldots \quad (33)$$

$$\epsilon = \cos Nh(e^{-Mh} + e^{Mh}) \quad \ldots \quad (34)$$

$$\kappa = \frac{e^{Mh}}{(M^2+N^2)}(N \sin Nh + M \cos Nh) \quad \ldots \quad (35)$$

$$\lambda = \frac{e^{Mh}}{(M^2+N^2)}(M \sin Nh - N \cos Nh) \quad \ldots \quad (36)$$

The transformed vertical velocity, $\bar{w}$, can be easily obtained by the integration of Eq. 17:

$$\bar{w} = \int_{-h}^{z}\frac{\partial \bar{w}}{\partial \xi}\,d\xi = -\int_{-h}^{z}\left(\frac{\partial \bar{u}}{\partial x} + \frac{\partial \bar{v}}{\partial y}\right)d\xi \quad \ldots \quad (37)$$

The resulting expression is lengthy; it appears as Eq. 60 in Appendix I.

Pressure Equation.—The boundary conditions of the vertical velocity lead to equations for the horizontal pressure distribution. Using $\bar{w} = 0$ at $z = 0$ in Eq. 60 leads to

$$a\frac{\partial^2 \bar{p}}{\partial x^2} + b\frac{\partial^2 \bar{p}}{\partial x \partial y} + c\frac{\partial^2 \bar{p}}{\partial y^2} + d\frac{\partial \bar{p}}{\partial x} + e\frac{\partial \bar{p}}{\partial y} = f \quad \ldots \quad (38)$$

in which the coefficients a, b, ..., f are functions of x and y.

Vertically Averaged Velocities.—A vertical average of the transformed velocities is

$$\bar{u} = \frac{1}{h}\int_{-h}^{0}\tilde{u}\,dz \quad \ldots \quad (39)$$

$$\bar{v} = \frac{1}{h}\int_{-h}^{0}\tilde{v}\,dz \quad \ldots \quad (40)$$

Eqs. 20 and 21 are substituted into Eqs. 39 and 40 and vertical integrations carried out to yield

$$\bar{u} = h_1\frac{\partial \bar{p}}{\partial x} + h_4\frac{\partial \bar{p}}{\partial y} + h_2\bar{\Gamma} + h_3\bar{\Delta} \quad \ldots \quad (41)$$

$$\bar{v} = -h_4\frac{\partial \bar{p}}{\partial x} + h_1\frac{\partial \bar{p}}{\partial y} + h_3\bar{\Gamma} - h_2\bar{\Delta} \quad \ldots \quad (42)$$

in which h_1, h_2, h_3, and h_4 are functions of x and y and are given in Appendix I. The condition that $\bar{u} = \bar{v} = 0$ at the edge of the lake is used to provide the boundary conditions on Eq. 38 by means of Eqs. 41 and 42.

Stream Function.—The numerical solution is actually carried out for the stream function defined as

$$\bar{u} = \frac{1}{h}\frac{\partial \bar{\psi}}{\partial y} \quad \ldots \quad (43)$$

$$\bar{v} = -\frac{1}{h}\frac{\partial \bar{\psi}}{\partial x} \quad \ldots \quad (44)$$

Thus, the integrated continuity equation

$$\int_{-h}^{0} \left(\frac{\partial \bar{u}}{\partial x} + \frac{\partial \bar{v}}{\partial y} \right) dz = \frac{\partial}{\partial x}(h\bar{u}) + \frac{\partial}{\partial y}(h\bar{v}) = 0 \quad \ldots \quad (45)$$

is automatically satisfied. Eqs. 43 and 44 are substituted into Eqs. 41 and 42, and the result solved for the pressure gradients:

$$\frac{\partial \bar{p}}{\partial x} = q_1 \frac{\partial \bar{\psi}}{\partial y} + q_2 \frac{\partial \bar{\psi}}{\partial x} + q_4 \bar{\Gamma} - q_3 \bar{\Delta} \quad \ldots \quad (46)$$

$$\frac{\partial \bar{p}}{\partial y} = q_2 \frac{\partial \bar{\psi}}{\partial y} - q_1 \frac{\partial \bar{\psi}}{\partial x} - q_3 \bar{\Gamma} - q_4 \bar{\Delta} \quad \ldots \quad (47)$$

The variables, q_1, q_2, q_3, q_4, are defined in Appendix I. By cross differentiation of Eqs. 46 and 47, the following equations is obtained for the stream function:

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + A(x, y; s) \frac{\partial \bar{\psi}}{\partial x} + B(x, y; s) \frac{\partial \bar{\psi}}{\partial y} + C(x, y; s) = 0 \quad \ldots \quad (48)$$

$$\text{in which} \quad A(x, y; s) = \frac{h}{h_1}(h_1^2 + h_4^2)\left(\frac{\partial q_1}{\partial x} + \frac{\partial q_2}{\partial y} \right) \quad \ldots \quad (49)$$

$$B(x, y; s) = \frac{h}{h_1}(h_1^2 + h_4^2)\left(\frac{\partial q_1}{\partial y} - \frac{\partial q_2}{\partial x} \right) \quad \ldots \quad (50)$$

$$C(x, y; s) = \frac{h}{h_1}(h_1^2 + h_4^2)\left[\frac{\partial}{\partial x}(q_3 \bar{\Gamma}) - \frac{\partial}{\partial y}(q_3 \bar{\Delta}) + \frac{\partial}{\partial y}(q_4 \bar{\Gamma}) + \frac{\partial}{\partial x}(q_4 \bar{\Delta}) \right] \quad \ldots \quad (51)$$

Eq. 48 is solved numerically under the condition that $\bar{\psi}$ is a constant along the boundary of the lake. Changing this constant (or making $\bar{\psi}$ variable over a small length) accounts for inflow and outflow. Once $\bar{\psi}$ is known, Eqs. 46 and 47 provide the pressure gradients; then, transformed point velocities are found from Eqs. 20, 21, and 60. The real point velocities are obtained from the inverse Laplace transform, which must be done numerically.

Finite Element Analysis

Eq. 48 is the same as Eq. 7 of Ref. 3, except that Eq. 48 contains the additional parameter, s. Exactly the same solution technique is employed as that described in Ref. 3. A separate solution must be performed for each value of the parameter. At this writing, a program using linear triangular elements has been modified for the transient calculation. However, a general program for the steady state is in use which provides a choice of linear triangular elements, cubic triangular elements, or quadratic isoparametric elements. Some mesh generation facility is also included. This latter program could be modified for the transie[illegible]alysis by simply changing some of the functions.

Laplace Transform Inversion

The procedure to this point has paralleled that of Refs. 3 and 6, with the addition of the transformed time derivative. The finite element program, with the final use of Eqs. 20, 21, and 60, gives the values of the transformed velocities. The inverse transform remains in order to find the physical velocities. The collocation method of Schapery (9) is used for the numerical transform inversion. The method assumes the functional form of the dependent variables with time in which the assumed function contains a number of undetermined coefficients. These coefficients are found from the solution of a set of simultaneous equations.

Following the lead of Shapery (9) the velocity is written as

$$u(t) = u_0 + \Delta u(t) \quad \ldots \quad (52)$$

with a similar expression for v and w. The u_0 is a constant and the last term is given by the finite Dirichlet series

$$\Delta u(t) = \sum_{i=1}^{n} a_i e^{-b_i t} \quad \ldots \quad (53)$$

Using Eq. 53, Eq. 52 is transformed and multiplied by s to yield

$$s\bar{u}(s) = u_0 + \sum_{j=1}^{n} \frac{a_j}{1 + \frac{b_j}{s}} \quad \ldots \quad (54)$$

The constant, u_0, is clearly the steady-state value of u for large time, or

$$u_0 = \lim_{s \to 0} s\bar{u}(s) \quad \ldots \quad (55)$$

If the n constants, b_j, are chosen by some means, then n values of s (e.g., s_i) can be selected so that Eq. 54 represents n equations in the undetermined coefficients, a_j. Schapery indicates that the error will be nearly minimized if the b_i are chosen equal to the selected s_i values, i.e.

$$b_i = s_i \quad j = 1, 2, \ldots, n \quad \ldots \quad (56)$$

Then n simultaneous algebraic equations occur in the a_j:

$$s_i \bar{u}(s_i) = u_0 + \sum_{j=1}^{n} \frac{a_j}{1 + \frac{s_j}{s_i}} \quad i = 1, 2, \ldots, n \quad \ldots \quad (57)$$

Choice of the s_i remains. First $s\bar{u}(s)$ versus log s is plotted as in Fig. 1. The significant range is that showing a definite variation of $s\bar{u}(s)$. Numerical experience (8,9) has shown that optimal results are achieved by selecting the s_i in a geometric sequence

$$\frac{s_{i-1}}{s_i} = r \quad \ldots \quad (58)$$

in which r is a fixed ratio. Thus, the upper and lower boun[illegible] the s_i are selected from the plot and the ratio, r, is fixed by choosing [illegible] vice-versa

The accuracy is generally increased by choosing a large n, but a too large n will result in numerical instability.

In the present study, the following values were used: $n = 8$, $r = 10$, $s_i = 2 \times 10^{-5}, 2 \times 10^{-4}, \ldots, 2 \times 10^{2}$. The resulting eight equations were solved using a standard Gauss-Jordan elimination procedure. The result yields a continuous solution for $u(t)$, $v(t)$, and $w(t)$. The procedure is repeated at each point in the x, y plane where the solution is desired. Thus, the number of constants, a_j, needed is the number of node points $\times$ n $\times$ 3. Although this number can

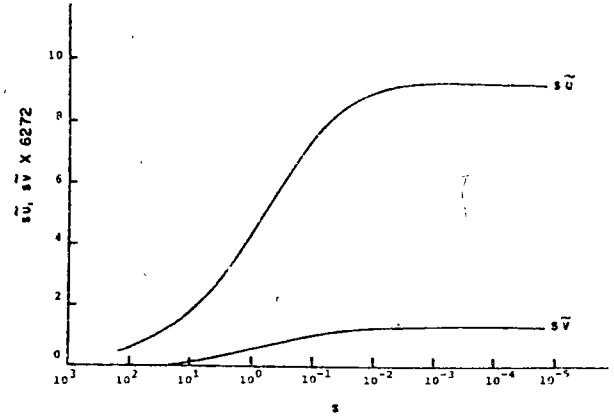


FIG. 1.—Transformed Velocity Components at Lake Surface for $x = 0.2$, $y = 0.5$ or $x = 0.8$, $y = 3.5$

be large (209 $\times$ 8 $\times$ 3 = 5,016 in the present study), the $n \times n$ coefficient matrix of the a_j needs to be inverted only once; i.e., Eq. 57 becomes

$$\frac{a_1}{1+\frac{s_1}{s_1}} + \frac{a_2}{1+\frac{s_2}{s_1}} + \ldots + \frac{a_n}{1+\frac{s_n}{s_1}} = s_1\tilde{u}(s_1) - u_0;$$

$$\frac{a_1}{1+\frac{s_1}{s_2}} + \frac{a_2}{1+\frac{s_2}{s_2}} + \ldots + \frac{a_n}{1+\frac{s_n}{s_2}} = s_2\tilde{u}(s_2) - u_0;$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\frac{a_1}{1+\frac{s_1}{s_n}} + \frac{a_2}{1+\frac{s_2}{s_n}} + \ldots + \frac{a_n}{1+\frac{s_n}{s_n}} = s_n\tilde{u}(s_n) - u_0 \quad \ldots\ldots\ldots\ldots\ldots (59)$$

in which only the right side changes with different positions in the x, y plane.

Numerical Results

The method was checked by solving a test case in an idealized basin. The problem chosen is that solved in Ref. 4, using a Fourier transform to remove

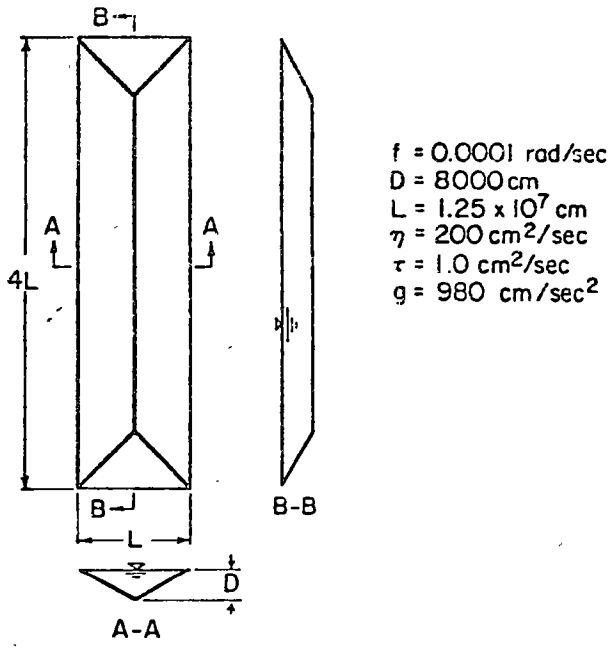


FIG. 2.—Rectangular Lake Configuration

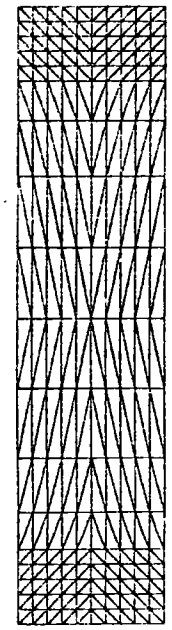


FIG. 3.—Finite Element Mesh Arrangement for Rectangular Lake

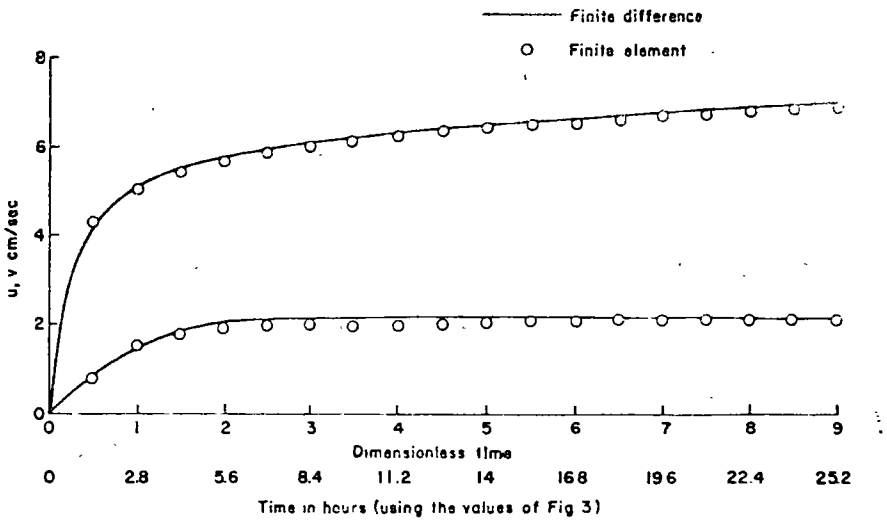


FIG 4.—Comparison of Two Different Types of Solution for Velocity Components at Lake Surface at Typical Point ($x = 0.3$, $y = 2.0$, $z = 0$)

vertical dependence and finite differences to step through time. The test basin is described in Ref. 4 and appears in Fig. 2 with the dimensional parameters.

The finite element grid is shown in Fig. 3. The grid consists of 360 elements with 209 nodal points. The finite element resolution is much greater than the finite difference resolution used in Ref. 4. Details of the finite element calculation are given in Ref. 3.

The results of the computation was compared to the results of the study of Ref. 4 at a large number of grid points. A typical comparison is shown in Fig. 4. The horizontal velocity components are shown at $x = 0.3$, $y = 2.0$, $z = 0$ (equal to the solution at $x = 0.7$, $y = 2.0$, $z = 0$ due to antisymmetry). Small differences between the two calculations are attributed to differences in resolution and the different method of handling the time variation. No statement can be made as to which is the most accurate. The differences appear sufficiently small to be of no practical importance.

Conclusions

The Laplace transform with numerical inversion technique appears to have a large efficiency advantage in the present problem, stemming especially from the fact that the vertical distribution of velocity can be stated explicitly in algebraic equations instead of in the rather awkward (for machine calculation) Fourier series or in a finite difference sense. The result is a velocity distribution continuous in depth and time and, using the automatic interpolation facility of the finite element formulation, continuous in the horizontal dimensions also.

Any three-dimensional time dependent problem is very large. This problem has been reduced to two dimensions plus an algebraic formula in depth plus a few (eight in the results presented herein) solutions to a "steady-state problem with parameter." These shortcuts have created a situation whereby the calculation is quick and economical and can be performed many times for parameter study. However, the presentation of results still remains a problem. It is not easy to picture the three velocity components varying with time in a three-dimensional space. Work is currently underway to develop computer graphics procedures which will resolve the presentation problem. The development of such procedures was, in fact, the major motivation for finding efficient solutions that could be stored in the computer and called at will for graphical presentation.

Acknowledgment

Work described in this paper was supported by the National Science Foundation under Research Grants GK23992 and ENG76-05029.

Appendix I.—Formulas

$$\bar{w} = \frac{1}{s^2 + 1}(s\nabla^2\bar{p})(z + h)$$

$$- \frac{1}{(M^2 + N^2)}\left(\frac{\partial c_3}{\partial y} - \frac{\partial c_4}{\partial x}\right)[e^{Mz}(M\cos Nz + N\sin Nz)$$

$$- e^{-Mh}(M\cos Nh - N\sin Nh) + e^{-Mz}(-M\cos Nz + N\sin Nz)$$

$$+ e^{Mh}(M\cos Nh + N\sin Nh)]$$

$$+ \frac{1}{(M^2 + N^2)}\left(\frac{\partial c_3}{\partial x} + \frac{\partial c_4}{\partial y}\right)[e^{Mz}(M\sin Nz - N\cos Nz)$$

$$+ e^{-Mh}(M\sin Nh + N\cos Nh) + e^{-Mz}(M\sin Nz + N\cos Nz)$$

$$+ e^{Mh}(M\sin Nh - N\cos Nh)]$$

$$- \frac{1}{(M^2 + N^2)^2}\left\{N\left(\frac{\partial\bar{\Gamma}}{\partial x} - \frac{\partial\bar{\Delta}}{\partial y}\right)[e^{Mz}(M\cos Nz + N\sin Nz)\right.$$

$$- e^{-Mh}(M\cos Nh - N\sin Nh)]$$

$$- M\left(\frac{\partial\bar{\Gamma}}{\partial x} - \frac{\partial\bar{\Delta}}{\partial y}\right)[e^{Mz}(M\sin Nz - N\cos Nz)$$

$$\left. - e^{-Mh}(-M\sin Nh - N\cos Nh)]\right\}$$

$$- \frac{1}{(M^2 + N^2)^2}\left\{M\left(\frac{\partial\bar{\Gamma}}{\partial y} + \frac{\partial\bar{\Delta}}{\partial x}\right)[e^{Mz}(M\cos Nz + N\sin Nz)\right.$$

$$- e^{-Mh}(M\cos Nh - N\sin Nh)]$$

$$+ N\left(\frac{\partial\bar{\Gamma}}{\partial y} + \frac{\partial\bar{\Delta}}{\partial x}\right)[e^{Mz}(M\sin Nz - N\cos Nz)$$

$$\left. - e^{-Mh}(-M\sin Nh - N\cos Nh)]\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (60)$$

$$h_1 = \frac{1}{\alpha h(s^2 + 1)}[-sh\alpha + (\gamma + \epsilon s)(\beta + \kappa) + (-\gamma s + \epsilon)(\delta + \lambda)] \quad \ldots\ldots \quad (61)$$

$$h_2 = \frac{1}{\alpha h}\left[\frac{\alpha}{M^2 + N^2}\left(\frac{2MN}{M^2 + N^2} + N\beta - M\delta\right) - (\gamma\beta + \epsilon\delta)(\beta + \kappa)\right.$$

$$\left. - (\beta\epsilon - \delta\gamma)(\delta + \lambda)\right] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (62)$$

$$h_3 = \frac{1}{\alpha h}\left[\frac{\alpha}{M^2 + N^2}\left(\frac{M^2 - N^2}{M^2 + N^2} + M\beta + N\delta\right) + (\beta\epsilon - \delta\gamma)(\beta + \kappa)\right.$$

$$\left. - (\beta\gamma + \delta\epsilon)(\delta + \lambda)\right] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (63)$$

$$h_4 = \frac{1}{\alpha h(s^2 + 1)}[-\alpha h + (\epsilon - \gamma s)(\beta + \kappa) - (\gamma + \epsilon s)(\delta + \lambda)] \quad \ldots\ldots\ldots \quad (64)$$

$$q_1 = \frac{h_1}{h(h_1^2 - h_4^2)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (65)$$

$$q_2 = \frac{h_4}{h(h_1^2 + h_4^2)} \quad \ldots \quad (66)$$

$$q_3 = \frac{h_1 h_3 + h_2 h_4}{h_1^2 + h_4^2} \quad \ldots \quad (67)$$

$$q_4 = \frac{h_3 h_4 - h_1 h_2}{h_1^2 + h_4^2} \quad \ldots \quad (68)$$

Appendix II.—References

1. Connor, J. J., and Wang, J., "Finite Element Modeling of Hydrodynamic Circulation," *Numerical Methods in Fluid Dynamics*, C. A. Brebbia and J. J. Connor, eds., Pentech Press, London, England, 1974, pp. 555-587.
2. Fried, I., "Finite-Element Analysis of Time-Dependent Phenomena," *American Institute of Aeronautics and Astronautics Journal*, Vol. 7, No. 6, June, 1969, pp. 1170-1173.
3. Gallagher, R. H., Liggett, J. A., and Chan, S. T. K., "Finite Element Shallow Lake Circulation Analysis," *Journal of the Hydraulics Division*, ASCE, Vol. 99, No. HY7, Proc. Paper 9855, July, 1973, pp. 1083-1096.
4. Liggett, J. A., "Unsteady Circulation in Shallow, Homogeneous Lakes," *Journal of the Hydraulics Division*, ASCE, Vol. 95, No. HY4, Proc. Paper 6686, July, 1969, pp. 1273-1288
5. Liggett, J. A., "A Cell Method for Computing Lake Circulation," *Journal of the Hydraulics Division*, ASCE. Vol. 96, No. HY3, Proc. Paper 7152, Mar., 1970, pp. 725-743.
6. Liggett, J. A., and Hadjitheodorou, C., "Circulation in Shallow Homogeneous Lakes," *Journal of the Hydraulics Division*, ASCE, Vol. 95, No. HY2, Proc. Paper 6454, Mar., 1969, pp. 609-620.
7. Lindh, G., and Bengtsson, L., "Wind-Induced Circulation in a Lake," *Bulletin Series A No. 10*, Division of Hydraulics, Institute of Technology, University of Lund, Lund, Sweden, 1972.
8. Rizzo, F. J., and Shippy, D. J., "A Method of Solution for Certain Problems of Transient Heat Conduction," *American Institute of Aeronautics and Astronautics Journal*, Vol. 8, No. 11, Nov., 1970, pp. 2004-2009.
9. Shapery, R. A., "Approximate Methods of Transform Inversion for Viscoelastic Stress Analysis," *Proceedings of the 4th U.S National Congress of Applied Mechanics*, Vol. 2, 1962, pp. 1075-1085.
10. Taylor, G., and Hood, P., "A Numerical Solution of the Navier Stokes Equations Using the Finite Element Technique," *Computers and Fluids*, Vol. 1, pp. 73-100.
11. Welander, P., "Wind Action on a Shallow Sea: Some Generalizations of Ekman's Theory," *Tellus*, Vol. 9, No. 1, Feb., 1957, pp. 45-52.

Appendix III.—Notation

The following symbols are used in this paper:

$A, B, C, c_1, c_2, c_3, c_4, a, b, c, d, e, f, h_1, h_2, h_3, h_4, q_1, q_2, q_3, q_4, \alpha, \beta, \gamma, \delta, \epsilon, \kappa, \lambda$ = function of x and y;
a_i, b_i, u_0, n = constants;
D = typical vertical dimension used to normalize depth;
f = Coriolis parameter;

g = acceleration of gravity;
h = normalized depth of lake;
L = typical horizontal dimension used to normalize width and length;
M = $\sqrt{2}\, m \sqrt{R} \cos(\phi/2)$;
m = $\sqrt{D^2 f/2\eta}$;
N = $\sqrt{2}\, m \sqrt{R} \sin(\phi/2)$;
p = local pressure;
R = $\sqrt{(1 + s^2)}$;
r = ratio of geometric series;
s = Laplace transform parameter;
t = time;
$\bar{u}, \bar{v}$ = average transformed velocity components in x and y directions, respectively;
x, y, z = Cartesian Coordinates with x and y in a horizontal plane and z positive upward and zero at the surface;
η = eddy viscosity;
ρ = fluid density;
τ_x, τ_y = surface wind stresses in x and y directions, respectively;
ϕ = $\tan^{-1}(1/s)$;
ψ = stream function; and
$\tilde{}$ = transformed variable.

Journal of the
HYDRAULICS DIVISION
Proceedings of the American Society of Civil Engineers

PROPERTIES OF CIRCULATION IN STRATIFIED LAKES

By James A. Liggett,[1] M. ASCE and Kwang K. Lee,[2] A. M. ASCE

INTRODUCTION

The complexity of circulation in a stratified body of water precludes an accurate analysis at the present time. It appears that such problems can be attacked by a large computing program (7); however, even a sophisticated computer program must contain a number of simplifying approximations. Moreover, the computer method has the disadvantage that the features of the circulation, the causes and effects, get lost in a massive program. The objective herein is to delineate features of motion in a stratified body of water in a way that explains observed phenomena and can be used to develop intuitive insight into the problem.

To accomplish this objective, many assumptions and approximations must be made. These approximations render the quantitative aspects of the results doubtful, although the writers believe that the quantitative results could be used in the absence of a more rigorous method. Only the problem of steady flow has been considered in this paper. Although the time response of a stratified lake is in doubt, it appears that a true steady flow seldom occurs. However, many of the features of a steady flow analysis are observed in lakes. The steady flow analysis can also be used to define an average condition; it is this average condition which is likely to be useful for design pur-

Note.—Discussion open until June 1, 1971. To extend the closing date one month, a written request must be filed with the Executive Director, ASCE. This paper is part of the copyrighted Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, Vol. 97, No. HY1, January, 1971. Manuscript was submitted for review for possible publication on February 19, 1970.

[1]Assoc. Prof., School of Civil Engrg., Cornell Univ., Ithaca, N.Y.
[2]Asst. Prof., College of Environmental Sciences, Univ. of Wisconsin, Green Bay, Wisc.; formerly Research Assoc., School of Civil Engrg., Cornell Univ., Ithaca, N.Y.

poses. The great mass of data of an unsteady analysis is often confusing and not useful; hence the climatology of a lake is better defined by averages.

PROBLEM FORMULATION

The summer stratification of a lake often consists of three layers: (1) the epilimnion, an upper layer of near-constant density; (2) the metalimnion, a middle layer of steep density gradient; and (3) the hypolimnion (1,6,7,13). The stratification is assumed to approximate a two-layer system herein. Fig. 1 indicates the idealization to a two-layer lake in which the epilimnion and hypolimnion are considered homogeneous.

The equations of motion are the linearized equations which have been used to describe the motion in homogeneous (8,9,10) and two-layer (2,7) lakes. The nonlinear terms are neglected due to a small Rossby number (i.e., the ratio

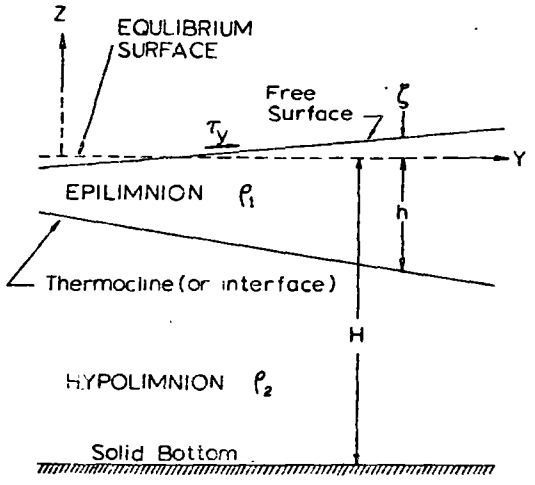


FIG. 1.—DEFINITION SKETCH OF LAKE

of inertial and rotational forces). Only vertical friction is considered; the Coriolis parameter is taken as a constant; and the pressure is assumed to be hydrostatic. The equations for each of the two layers are:

$$-fv_k = -\frac{1}{\rho_k}\frac{\partial p_k}{\partial x} + \eta_k \frac{\partial^2 u_k}{\partial z^2} \quad \ldots\ldots (1)$$

$$fu_k = -\frac{1}{\rho_k}\frac{\partial p_k}{\partial y} + \eta_k \frac{\partial^2 v_k}{\partial z^2} \quad \ldots\ldots (2)$$

$$g = -\frac{1}{\rho_k}\frac{\partial p_k}{\partial z} \quad \ldots\ldots (3)$$

$$\frac{\partial u_k}{\partial x} + \frac{\partial v_k}{\partial y} + \frac{\partial w_k}{\partial z} = 0 \quad \ldots\ldots (4)$$

in which the subscript $k = 1$ for the epilimnion and $k = 2$ for the hypolimnion.

The coordinates are x, y, and z with x-y plane horizontal and z vertically upward; the velocities corresponding to the x, y, z directions are u, v, w respectively; f = Coriolis parameter; ρ = density; p = pressure; η = kinematic eddy viscosity; and g = acceleration of gravity.

The boundary conditions on the epilimnion are:

$$\tau_x = \eta_1\rho_1 \frac{\partial u_1}{\partial z}, \quad \tau_y = \eta_1\rho_1 \frac{\partial v_1}{\partial z} \quad \text{at } z = 0 \quad \ldots\ldots (5)$$

$$\rho_1\eta_1 \frac{\partial u_1}{\partial z} = G_x, \quad \rho_1\eta_1 \frac{\partial v_1}{\partial z} = G_y \quad \text{at } z = -h \quad \ldots\ldots (6)$$

in which τ_x and τ_y = components of surface wind stress (assumed as known) and G_x and G_y = shear stresses acting on the epilimnion at the thermocline. The primary (steady state) driving forces for currents in the hypolimnion are the pressure gradient and the interface stresses, G_x and G_y. Currents in the epilimnion are caused by the wind stresses, τ_x and τ_y, and are modified by the pressure gradient and interface stresses.

METHOD OF SOLUTION

From the preceding equations, the velocity components in the upper layer are found to be

$$u_1 = -\frac{1}{f\rho_1}\frac{\partial p_1}{\partial y} + \cos qz\,[A_3 \exp(qz) - A_4 \exp(-qz)]$$
$$- \sin qz\,[A_1 \exp(qz) - A_2 \exp(-qz)] \quad \ldots\ldots (7)$$

$$v_1 = \frac{1}{f\rho_1}\frac{\partial p_1}{\partial x} + \cos qz\,[A_1 \exp(qz) + A_2 \exp(-qz)]$$
$$+ \sin qz\,[A_3 \exp(qz) + A_4 \exp(-qz)] \quad \ldots\ldots (8)$$

in which $q = \sqrt{f/2\eta_1}$. The values of A_1, A_2, A_3 and A_4 are determined by the boundary conditions. As a shorthand notation, define $\Delta = -2(\cosh 2qh - \cos 2qh)$; $T_1 = (\tau_y - \tau_x)/2q\eta_1\rho_1$; $T_2 = (\tau_y + \tau_x)/2q\eta_1\rho_1$; $T_3 = (G_x + G_y)/2q\eta_1\rho_1$; and $T_4 = (G_y - G_x)/2q\eta_1\rho_1$.

The A's are: $A_1 = \{T_1[\cos 2qh - \exp(2qh)] + T_2(\sin 2qh)$

$$- T_3(2\sin qh \cosh qh) + T_4(2\cos qh \sinh qh)\}/\Delta \quad \ldots\ldots (9)$$

$$A_2 = \{T_1[\exp(-2qh) - \cos 2qh] + T_2(\sin 2qh)$$
$$- T_3(2\sin qh \cosh qh) + T_4(2\cos qh \sinh qh)\}/\Delta \quad \ldots\ldots (10)$$

$$A_3 = \{T_1(-\sin 2qh) - T_2[\exp(2qh) - \cos 2qh]$$
$$+ T_3(2\cos qh \sinh qh) + T_4(2\sin qh \cosh qh)\}/\Delta \quad \ldots\ldots (11)$$

$$A_4 = \{T_1(\sin 2qh) - T_2[\exp(-2qh) - \cos 2qh]$$
$$- T_3(2\cos qh \sinh qh) - T_4(2\sin qh \cosh qh)\}/\Delta \quad \ldots\ldots (12)$$

From Eqs. 7 and 8, the epilimnion velocities are functions of the pressure gradient, wind shear, and interface stress. Approximations are found in the following sections for the interface shear and pressure gradients.

Interface Shear.—Let $W_k = u_k + iv_k$, in which W = complex horizontal

velocity and $i = \sqrt{-1}$. Omitting subscript k, Eqs. 1 and 2 become:

$$fW = \frac{1}{\rho}\left(i\frac{\partial p}{\partial x} - \frac{\partial p}{\partial y}\right) - i\eta\frac{\partial^2 W}{\partial z^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots : (13)$$

The horizontal velocity is separated into two parts as $W = W' + W_g$; in which $W_g = (1/\rho f)\,[i\,(\partial p/\partial x) - \partial p/\partial y]$; and thus

$$u_g = -\frac{1}{\rho f}\frac{\partial p}{\partial y},\quad v_g = \frac{1}{\rho f}\frac{\partial p}{\partial x} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (14)$$

are the geostrophic velocities. Then

$$fW' = -i\eta\frac{\partial^2 W'}{\partial z^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (15)$$

in which it is assumed, through the hydrostatic approximation that, $\partial^2 W_g/\partial z^2 = 0$. Solving Eq. 15 and adding the geostrophic component yields

$$W_k = \hat{A}_k \exp\left[\left(i\frac{f}{\eta_k}\right)^{1/2} z\right] + \hat{B}_k \exp\left[-\left(i\frac{f}{\eta_k}\right)^{1/2} z\right] + W_{gk} \quad (16)$$

Constants $\hat{A}_k$ and $\hat{B}_k$ could be determined from the free surface and bottom boundary conditions as well as from the interface conditions; however, the result would be rather complex. The purpose of this exercise is to derive an approximate expression for the interface shear; therefore, it is assumed that the solution near the interface is not greatly changed by ignoring the boundary conditions on the free surface and the bottom. Obviously such an approximation is valid if both of the two layers are thick, so that motion near the thermocline is determined primarily by pressure gradients (geostrophic flow).

Eq. 16 would have resulted had the origin of the coordinate system been at the thermocline. Let $z' = z + h$; then

$$W_k = A'_k \exp\left[\left(i\frac{f}{\eta_k}\right)^{1/2} z'\right] + B'_k \exp\left[-\left(i\frac{f}{\eta_k}\right)^{1/2} z'\right] + W_{gk} \quad (17)$$

If the free surface and lower boundary are far away, then

$$W_1 = B'_1 \exp\left[-\left(i\frac{f}{\eta_1}\right)^{1/2} z'\right] + W_{1g} \quad \ldots\ldots\ldots\ldots\ldots (18)$$

$$W_2 = A'_2 \exp\left[\left(i\frac{f}{\eta_2}\right)^{1/2} z'\right] + W_{2g} \quad \ldots\ldots\ldots\ldots\ldots (19)$$

approximately. The appropriate interface conditions are:

$$W_1 = W_2 \text{ and } \rho_1\eta_1\frac{\partial W_1}{\partial z'} = \rho_2\eta_2\frac{\partial W_2}{\partial z'} \text{ at } z' = 0 \quad \ldots\ldots\ldots\ldots (20)$$

Therefore $B'_1 = -\dfrac{\rho_2\eta_2^{1/2}}{\rho_1\eta_1^{1/2} + \rho_2\eta_2^{1/2}}(W_{1g} - W_{2g})$ . . . . . (21)

$$A'_2 = \frac{\rho_1\eta_1^{1/2}}{\rho_1\eta_1^{1/2} + \rho_2\eta_2^{1/2}}(W_{1g} - W_{2g}) \quad \ldots\ldots\ldots\ldots\ldots (22)$$

The interface stresses are found to be

$$G_x = K(u_{1g} - u_{2g} - v_{1g} + v_{2g}) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (23)$$

$$G_y = K(u_{1g} - u_{2g} + v_{1g} - v_{2g}) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (24)$$

in which $K = \sqrt{\dfrac{f\eta_1\eta_2}{2}}\left(\dfrac{\rho_1\rho_2}{\rho_1\eta_1^{1/2} + \rho_2\eta_2^{1/2}}\right)$ . . . . . (25)

From the hydrostatic condition (Eq. 3) and the condition that the pressure is continuous at the interface ($p_1 = p_2$ at $z' = 0$)

$$\frac{\partial p_1}{\partial x} = \rho_1 g\frac{\partial \zeta}{\partial x} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (26)$$

$$\frac{\partial p_1}{\partial y} = \rho_1 g\frac{\partial \zeta}{\partial y} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (27)$$

$$\frac{\partial p_2}{\partial x} = \rho_1 g\frac{\partial \zeta}{\partial x} + \rho_1 g\frac{\partial h}{\partial x} - \rho_2 g\frac{\partial h}{\partial x} \quad \ldots\ldots\ldots\ldots\ldots (28)$$

$$\frac{\partial p_2}{\partial y} = \rho_1 g\frac{\partial \zeta}{\partial y} + \rho_1 g\frac{\partial h}{\partial y} - \rho_2 g\frac{\partial h}{\partial y} \quad \ldots\ldots\ldots\ldots\ldots (29)$$

Using the definition of geostrophic flow (Eq. 14), the interface stresses can be written in terms of the surface and interface gradients as:

$$G_x = -\frac{gK}{f}\left(1 - \frac{\rho_1}{\rho_2}\right)\left(\frac{\partial \zeta}{\partial x} + \frac{\partial \zeta}{\partial y} + \frac{\partial h}{\partial x} + \frac{\partial h}{\partial y}\right) \quad \ldots\ldots\ldots\ldots (30)$$

$$G_y = \frac{gK}{f}\left(1 - \frac{\rho_1}{\rho_2}\right)\left(\frac{\partial \zeta}{\partial x} - \frac{\partial \zeta}{\partial y} + \frac{\partial h}{\partial x} - \frac{\partial h}{\partial y}\right) \quad \ldots\ldots\ldots\ldots (31)$$

Horizontal Velocities.—Eqs. 7 and 8 are derived for a wind stress in any direction. For the present research, they can be simplified without loss of generality by considering only a wind stress in the y-direction, i.e., $\tau_x = 0$. Substituting Eqs. 26 through 31 into Eqs. 7 and 8, the current velocity under a y-shear on the free surface is

$$u_1 = -\frac{g}{f}\frac{\partial \zeta}{\partial y} + \frac{1}{q\eta_1\rho_1}\left\{\frac{\tau_y}{2}\right.\{\cosh qz\,\sin q(2h + z)$$

$$- \sin qz\,\cosh q(2h + z) + \cos qz\,\sinh q(2h + z)$$

$$- \sinh qz\,\cos q(2h + z)] + \left(\frac{2Kg}{f}\right)\left(1 - \frac{\rho_1}{\rho_2}\right)\left[\left(\frac{\partial \zeta}{\partial y}\right.\right.$$

$$\left.+ \frac{\partial h}{\partial y}\right)(\sinh qh\,\cos qh\,\cos qz\,\cosh qz$$

$$+ \sin qh\,\cosh qh\,\sinh qz\,\sin qz)$$

$$-\left(\frac{\partial \zeta}{\partial x} + \frac{\partial h}{\partial x}\right)(\sin qh\,\cosh qh\,\cosh qz\,\cos qz$$

$$\left.\left.- \cos qh\,\sinh qh\,\sin qz\,\sinh qz)\right]\right\}/(\cosh 2qh - \cos 2qh) \quad \ldots (32)$$

$$v_1 = \frac{g}{f}\frac{\partial \zeta}{\partial x} - \frac{1}{q\eta_1\rho_1}\left\{\frac{\tau_y}{2}\left[\sinh qz \cos q(2h + z)\right.\right.$$

$$- \cos qz \sinh q(2h + z) + \cosh qz \sin q(2h + z)$$

$$\left. - \sin qz \cosh q(2h + z)\right] + \left(\frac{2Kg}{f}\right)\left(1 - \frac{\rho_1}{\rho_2}\right)\left[\left(\frac{\partial \zeta}{\partial y}\right.\right.$$

$$+ \frac{\partial h}{\partial y}\Big)(\cos qz \sin qh \cosh qh \cosh qz$$

$$- \cos qh \sin qz \sinh qh \sinh qz)$$

$$+ \left(\frac{\partial \zeta}{\partial x} + \frac{\partial h}{\partial x}\right)(\cos qz \cos qh \sinh qh \cosh qz$$

$$\left.\left. + \sin qz \sin qh \cosh qh \sinh qz)\right]\right\}/(\cosh 2qh - \cos 2qh) \quad \ldots\ (33)$$

Surface and Thermocline Slopes.—The surface and thermocline slopes create pressure gradients which extend to the bottom of the lake. These slopes are the remaining unknown factors in the velocity equations. The primary cause of the pressure gradients is surface wind stress, although the pressure gradients are modified by the Coriolis force and the bottom stress. The latter effects will be neglected in this approximate theory.

Integrating Eqs. 1 and 2 with $k = 1, 2$ and with the wind stress in the y-direction ($\tau_x = 0$) yields

$$\int_{-h}^{\zeta} \frac{\partial p_1}{\partial x}\,dz - \eta_1\rho_1 \int_{-h}^{\zeta} \frac{\partial^2 u_1}{\partial z^2}\,dz = \frac{\partial p_1}{\partial x}(\zeta + h) + G_x = 0 \quad \ldots\ (34)$$

$$\int_{-h}^{\zeta} \frac{\partial v_1}{\partial y}\,dz - \eta_1\rho_1 \int_{-h}^{\zeta} \frac{\partial^2 v_1}{\partial z^2}\,dz = \frac{\partial p_1}{\partial y}(\zeta + h)$$

$$- (\tau_y - G_y) = 0 \quad \ldots\ (35)$$

$$\int_{-H}^{-h} \frac{\partial p_2}{\partial x}\,dz - \eta_2\rho_2 \int_{-H}^{-h} \frac{\partial^2 u_2}{\partial z^2}\,dz = \frac{\partial p_2}{\partial x}(-h + H) - G_x = 0 \quad (36)$$

$$\int_{-H}^{-h} \frac{\partial p_2}{\partial y}\,dz - \eta_2\rho_2 \int_{-H}^{-h} \frac{\partial^2 v_2}{\partial z^2}\,dz = \frac{\partial p_2}{\partial y}(-h + H) - G_y = 0 \quad \ldots\ (37)$$

Using Eqs. 26 through 29, Eqs. 34 through 37 become

$$\rho_1(\zeta + h)\frac{\partial \zeta}{\partial x} - \frac{K}{f}\left(1 - \frac{\rho_1}{\rho_2}\right)\left(\frac{\partial h}{\partial x} + \frac{\partial h}{\partial y}\right) = 0 \quad \ldots\ (38)$$

$$\rho_1(\zeta + h)\frac{\partial \zeta}{\partial y} + \frac{K}{f}\left(1 - \frac{\rho_1}{\rho_2}\right)\left(\frac{\partial h}{\partial x} - \frac{\partial h}{\partial y}\right) = \frac{\tau_y}{g} \quad \ldots\ (39)$$

$$\rho_2(H - h)\left[\left(1 - \frac{\rho_1}{\rho_2}\right)\frac{\partial h}{\partial x} - \frac{\rho_1}{\rho_2}\frac{\partial \zeta}{\partial x}\right]$$

$$- \frac{K}{f}\left(1 - \frac{\rho_1}{\rho_2}\right)\left(\frac{\partial h}{\partial x} + \frac{\partial h}{\partial y}\right) = 0 \quad \ldots\ (40)$$

$$\rho_2(H - h)\left[\left(1 - \frac{\rho_1}{\rho_2}\right)\frac{\partial h}{\partial y} - \frac{\rho_1}{\rho_2}\frac{\partial \zeta}{\partial y}\right]$$

$$+ \frac{K}{f}\left(1 - \frac{\rho_1}{\rho_2}\right)\left(\frac{\partial h}{\partial x} - \frac{\partial h}{\partial y}\right) = 0 \quad \ldots\ (41)$$

In Eqs. 38 through 41 $\partial\zeta/\partial x$ has been neglected when added to $\partial h/\partial x$, since $\partial\zeta/\partial x$ is of the order $[1 - (\rho_2/\rho_1)]\,\partial h/\partial x$, and $\partial\zeta/\partial y$ has been neglected when added to $\partial h/\partial y$. Solving Eqs. 38 through 41 for the four unknown gradients yields:

$$\frac{\partial h}{\partial y} = \frac{\tau_y}{g(\rho_2 - \rho_1)D_E}\,\frac{1 - \alpha}{(1 - \alpha)^2 + \alpha^2} \quad \ldots\ (42)$$

$$\frac{\partial h}{\partial x} = \frac{\tau_y}{g(\rho_2 - \rho_1)D_E}\,\frac{\alpha}{(1 - \alpha)^2 + \alpha^2} \quad \ldots\ (43)$$

$$\frac{\partial \zeta}{\partial y} = \frac{\tau_y}{g\rho_1 D_E}\,\frac{(1 - \alpha)(D_H - 2\alpha D_E) + D_E}{(D_E + D_H)\left[(1 - \alpha)^2 + \alpha^2\right]} \quad \ldots\ (44)$$

$$\frac{\partial \zeta}{\partial x} = \frac{\tau_y}{\rho_1 g D_E}\,\frac{D_H}{D_E + D_H}\,\frac{\alpha}{(1 - \alpha)^2 + \alpha^2} \quad \ldots\ (45)$$

in which $D_E = h + \zeta$ = epilimnion depth; $D_H = H - h$ = hypolimnion depth; and $\alpha = (K/\rho_2 f D_E)\left[1 + (D_E/D_H)\right]$.

Due to the simplifications made in obtaining Eqs. 18 and 19 from Eq. 17, the above equations are valid only for sufficiently thick epilimnion and hypolimnion. The writers believe that $\alpha < 0.3$ is an appropriate limit on epilimnion thickness, since $\partial h/\partial y = 0$ at $\alpha = 1$ and $\partial(\partial h/\partial y)/\partial\alpha = 0$ at $\alpha = 0.297$. In the following discussion $\alpha < 0.3$ is assumed.

Obviously, all the derivatives of h and ζ are positive, indicating that the free surface slopes upward in the downwind direction and to the right of the wind. The interface slopes downward in the downwind direction and to the right of the wind. Since α is small $\partial h/\partial y > \partial h/\partial x$, indicating that the slope of the interface along the direction of the wind is greater than crosswind.

Eqs. 44 and 45 are simplified if the hypolimnion thickness is large, $D_H \to \infty$, to

$$\frac{\partial \zeta}{\partial y} = \frac{\tau_y}{g\rho_1 D_E}\,\frac{1 - \alpha}{(1 - \alpha)^2 + \alpha^2}\ ;\ \frac{\partial \zeta}{\partial x} = \frac{\tau_y}{g\rho_1 D_E}\,\frac{\alpha}{(1 - \alpha)^2 + \alpha^2} \quad (46)$$

Eq. 46 look much like Eqs. 42 and 43 except that the effective density is ρ_1 instead of $\rho_2 - \rho_1$. Thus the interface will tend to slope by a much greater amount:

$$\frac{\partial h}{\partial y} \approx \frac{\rho_1}{\rho_2 - \rho_1}\frac{\partial \zeta}{\partial y}\ ;\ \frac{\partial h}{\partial x} \approx \frac{\rho_1}{\rho_2 - \rho_1}\frac{\partial \zeta}{\partial x} \quad \ldots\ (47)$$

than the free surface.

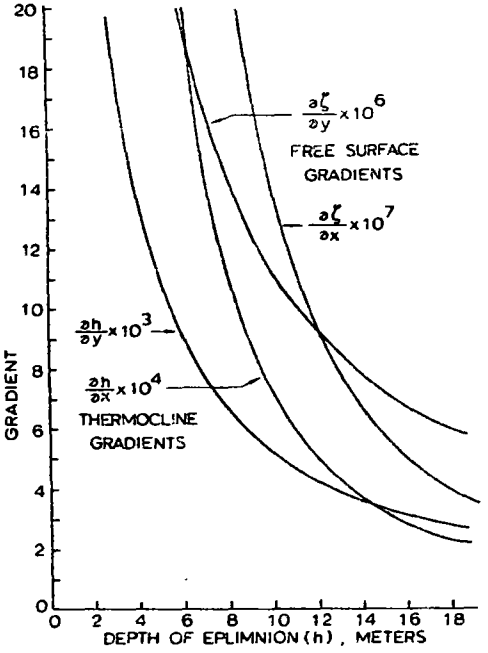


FIG. 2.—SURFACE AND THERMOCLINE GRADIENTS VERSUS DEPTH OF EPILIMNION

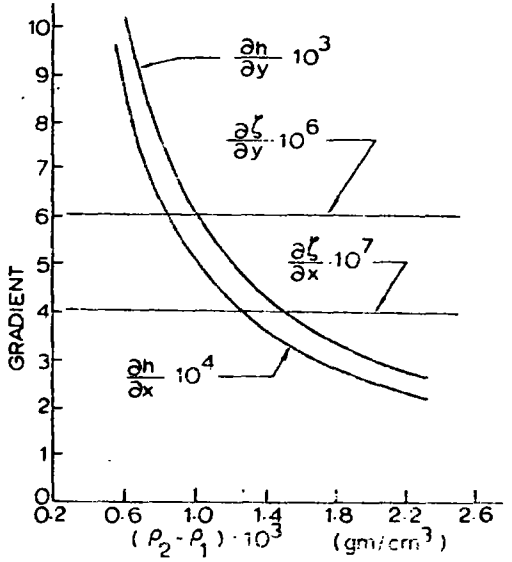


FIG. 3.—EFFECT OF DENSITY DIFFERENCE ON GRADIENTS OF SURFACE AND THERMOCLINE

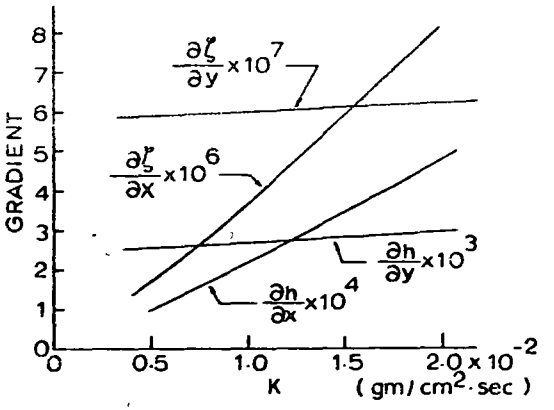


FIG. 4.—EFFECT OF EDDY VISCOSITY ON GRADIENTS OF SURFACE AND THERMOCLINE (SEE EQS. 23 and 24 FOR DEFINITION OF K)

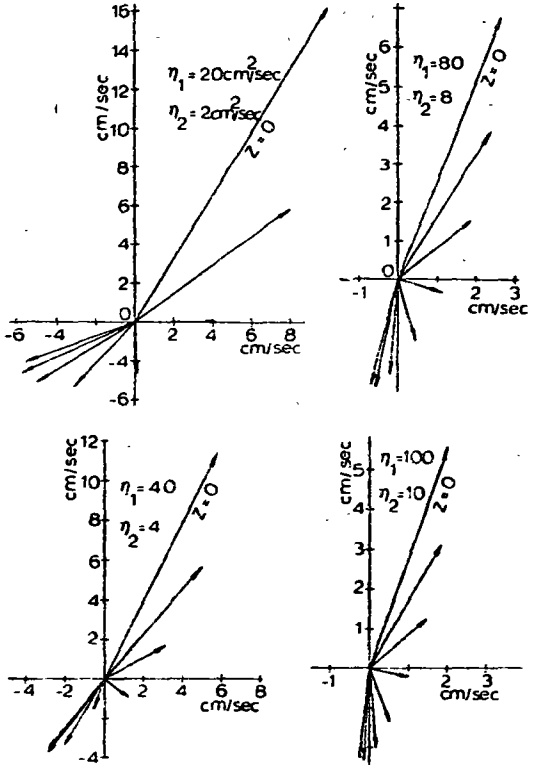


FIG. 5.—EPILIMNION HORIZONTAL VELOCITIES FOR DIFFERENT EDDY VISCOSITIES (ARROWS REPRESENT VELOCITY VECTORS AT DEPTHS $z = 0$; $z = -h/7$, $z = -2h/7$, . . . , $z = -6h/7$, $z = -h$.)

In the Great Lakes the cold hypolimnion water occasionally touches the surface in a cold water upwelling (2,3,4,7). Usually such upwelling occurs near the beach during an offshore wind. However, it may also occur if the wind is blowing parallel to the shore from the right (facing the lake) to the left (2,3,11).

NUMERICAL RESULTS

Most of the results can be more easily discussed if reasonable numerical values are assigned to the parameters. The following values will be used for

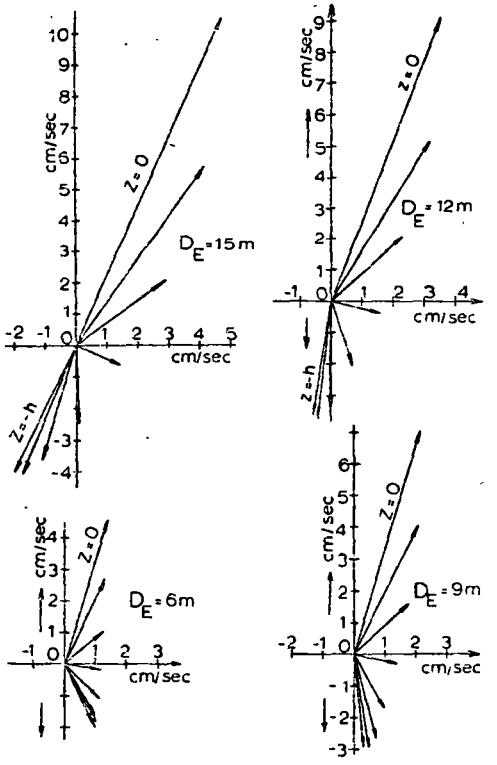


FIG 6.—EPILIMNION HORIZONTAL VELOCITIES FOR DIFFERENT EPILIMNION THICKNESS (ARROWS REPRESENT VELOCITY VECTORS AT DEPTHS $z = 0$; $z = -2h/7$; . . . ; $z = -6h/7$; $z = -h$)

the calculations in this paper unless otherwise noted: $f = 10^{-4}$ per sec; $g = 980$ cm per sec; $\tau_x = 0$; $\tau_y = 1$ dyne per sq cm; $\rho_1 = 0.99777$ gm per cu cm; $\rho_2 = 0.99997$ gm per cu cm; $\eta_1 = 40$ sq cm per sec; $\eta_2 = 4$ sq cm per sec; $H = 80$ m; and $h = 18$ m. The values are of the order of magnitude of those found or assumed by others for the Great Lakes (2,5,7,9).

Fig. 2 indicates the slopes of the free surface and thermocline using the

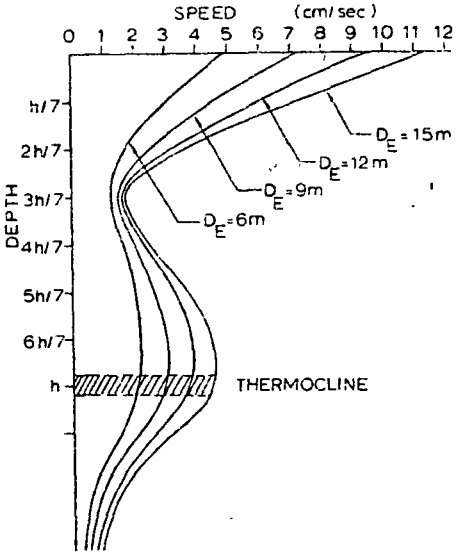


FIG. 7.—CURRENT SPEED VERSUS DEPTH DIAGRAM FOR DIFFERENT THICKNESSES OF EPILIMNION

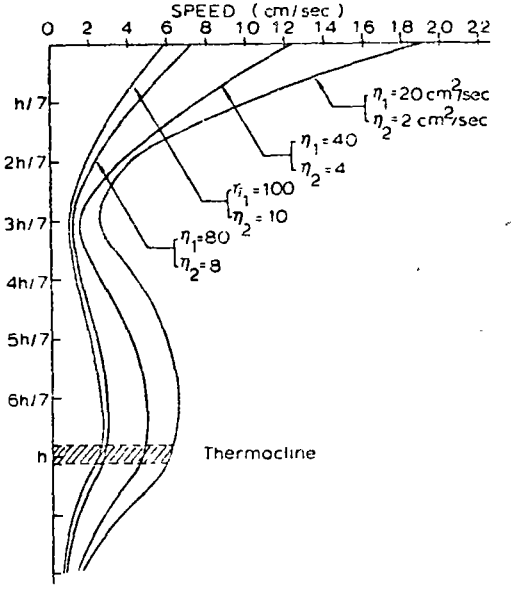


FIG. 8.—CURRENT SPEED VERSUS DEPTH DIAGRAM FOR DIFFERENT EDDY VISCOSITIES

above numbers, except that the epilimnion depth is allowed to vary. The strongest gradient is in the thermocline, which is tilted downward in the downwind direction. The thermocline is also tilted downward to the right of the wind, but the crosswind tilt is an order of magnitude less than the downwind tilt. The free surface is tilted in the opposite direction from the thermocline with the crosswind tilt again an order of magnitude less than the downwind tilt.

Fig. 3 indicates the tilting of the surface and thermocline as a function of the density difference. As expected, the density difference has a large influence in the thermocline tilt, but practically no influence in the free surface tilt. When the density difference is small, the thermocline may tilt very steeply so as to be unstable. Such a mechanism seems to cause the fall over-

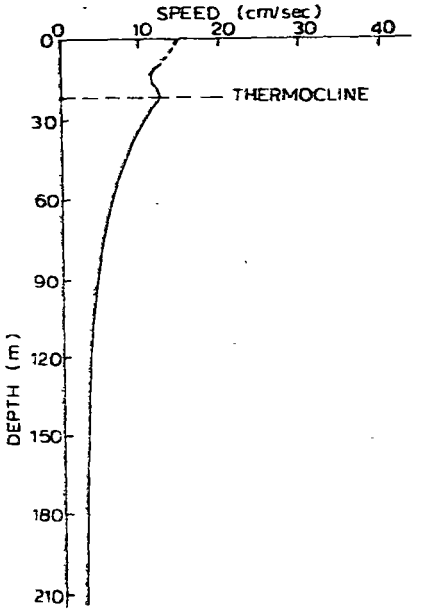


FIG. 9.—AVERAGE SUMMER CURRENT SPEEDS IN LAKE MICHIGAN (AFTER J. L. VERBER, Ref. 14)

turn when the epilimnion cools to such a degree that a moderate wind can upset the natural stability of the stratification.

Fig. 4 shows the tilt as a function of the parameter K. Eqs. 32 and 33 indicate that the velocities are complex functions of the various parameters even when effects of sideboundaries and bottom topography are excluded. The velocities are linearly proportional to wind stress. Thus an increase in wind stress only increases the velocities in magnitude without a directional change, neglecting the effect a higher wind stress might have on the eddy viscosity.

Fig. 5 indicates effect of epilimnion thickness on epilimnion velocities. The thicker epilimnion has greater velocities for the same wind stress and has an Ekman spiral which is more nearly complete.

Fig. 6 indicates the effect of eddy viscosities on epilimnion velocities. The eddy viscosities are largely unknown; thus it is important to determine the error made in current velocity due to incorrectly guessing eddy viscosity. The smaller eddy viscosities are associated with large velocities and more complete spirals.

Fig. 7 shows current speed as a function of depth for different epilimnion thicknesses. Fig. 8 shows current speed as a function of depth for different eddy viscosities. The average summer current speed as determined by Verber (14) is shown in Fig. 9. Eqs. 32 and 33 display the interaction within the stratified lake due to the different effects of each factor with respect to depth. Characteristically, the speed is maximum at the surface, decreasing to a relative minimum at three-sevenths the epilimnion depth, then increasing to a relative maximum just above the thermocline.

CONCLUSIONS

The present study is an attempt to gain some insight into the complex problem of circulation in a stratified lake. If the effect of the lateral boundaries and the bottom topography are neglected, reasonably simple equations can be derived for the slopes of the free surface and the thermocline and for the circulation. The approximate results present a satisfactory explanation for some of the observed phenomena. The results also indicate the effects of varying parameters of the problem as indicated in Figs. 2 through 8.

The quantative accuracy of this study is doubtful because of the numerous approximations. However, the results are in approximate quantitative agreement with the computer calculations of Lee and Liggett (7) in the region where the effects of the bottom boundary condition and the shore configuration become less important.

The response time of a stratified lake is unknown, but it is probably long, of the order of several days. In this respect the steady-state circulation may rarely, if ever, be attained in a natural lake. However, the steady-state condition might be expected to represent an average condition under average winds. Evidently, it is this average which would be most useful for purposes of design, for obtaining flushing data, for calculating the transport of physical quantities, etc. Since the problem appears to be linear to a good approximation (certainly to better accuracy than many of the approximations contained herein), the effects of through flow, or other non-wind-driven circulation, may be added.

The writers hope that the engineer can combine this type of study with field data (e.g., velocity measurements or drift card data) to infer the general circulation in any given lake.

APPENDIX I.—REFERENCES

1. Church, P. E., "The Annual Temperature Cycle of Lake Michigan I. Cooling from Late Autumn to the Terminate Point 1941-42," Institute of Meteorology, University of Chicago, *Miscellaneous Reports*, No. 4, [illegible]

2. Csanady, G. T., "Wind Driven Summer Circulation in the Great Lakes," *Journal of Geophysical Research*, Vol. 73, No. 8, April, 1968.
3. Csanady, G. T., and Ellenton, H. K., *Douglas Point Report*, Great Lakes Institute, Univ. of Toronto, PR 12 1962.
4. Hale, A. M., "Internal Waves in Lake Huron," *Hydrodynamics Studies on Lake Huron at Baie du Doré*, Great Lakes Institute, Univ. of Toronto, PR 19, 127, 1964.
5. Harleman, D. R. F., et al., "The Feasibility of a Dynamic Model of Lake Michigan," Pub. No. 9, Great Lakes Research Division, Univ. of Michigan, 1962.
6. Henson, E. B., Bradshaw, A. S., and Chandler, D. C., "The Physical Limnology of Cayuga Lake, New York," *Memoir* 378, Agriculture Experiment Station, New York State College of Agriculture, Cornell Univ., Ithaca, N.Y., Aug., 1961.
7. Lee, K. K., and Liggett, J. A., "Computation of Wind Driven Circulation in Stratified Lakes," *Journal of the Hydraulics Division*, ASCE, Vol. 96, No. HY10, Proc. Paper 7634, Oct., 1970.
8. Liggett, J. A., "A Cell Method for Computing Lake Circulation," *Journal of the Hydraulics Division*, ASCE, Vol. 96, No. HY3, Proc. Paper 7152, March, 1970, pp. 725–743.
9. Liggett, J. A., "Unsteady Circulation in Shallow Lakes," *Journal of the Hydraulics Division*, ASCE, Vol. 95, No. HY4, Proc. Paper 6686, July, 1969, pp. 1273–1288.
10. Liggett, J. A., and Hadjitheodorou, C., "Circulation in Shallow Homogenous Lakes," *Journal of the Hydraulics Division*, ASCE, Vol. 95, No. HY2, Proc. Paper 6454, March 1969, pp. 609–620.
11. Moffet, J., "An Instance of Upwelling along the East Shore of Lake Michigan, 1955," *Proceedings*, 5th Conference on Great Lakes Research, 1962.
12. Platzman, G. W., "The Dynamical Prediction of Wind Tides on Lake Erie," *Meteorological Monographs*, American Meteorological Society, Vol. 4, No. 26, 1963.
13. Rogers, G. K., and Anderson, D. V., "The Thermal Structure of Lake Ontario," Pub. No. 10, Great Lake Research Division, Univ. of Michigan, 1963.
14. Verber, J. L., "Current Profiles to Depth in Lake Michigan," Pub. No. 13, Great Lakes Research Division, Univ. of Michigan, 1965.

APPENDIX II.—NOTATION

The following symbols are used in this paper:

$A_1, A_2, A_3, A_4, \hat{A}, \hat{B}, A', B'$ = functions of x and y;
D_E = depth of epilimnion;
D_H = depth of hypolimnion;
f = Coriolis parameter;
g = acceleration of gravity;
G_x, G_y = components of interface stress at thermocline;
h = position function of epilimnion;
H = position function of lake bottom;
$i = \sqrt{-1}$;
k = subscript ($k = 1$ for epilimnion; $k = 2$ for hypolimnion);
K = constant (see Eqs. 23 and 24);
p = pressure;
q = constant;
T_1, T_2, T_3, T_4 = constant functions;
u, v, w = velocity components in x, y and z directions;

u_g, v_g = horizontal geostrophic velocity components;
W = complex horizontal velocity;
W' = transformed complex horizontal velocity;
W_g = complex geostrophic velocity;
x, y, z = three-dimensional coordinates;
x', y', z' = transformed coordinates;
α = dimensionless constant;
Δ = function of x and y;
ζ = position function of free surface;
η = eddy viscosity;
ρ = density of fluid; and
τ_x, τ_y = wind stress components in x and y directions.

Computers & Fluids, Vol. 1, pp. 119–132. Pergamon Press, 1973. Printed in Great Britain

HIGHER-ORDER FINITE ELEMENT ANALYSIS OF LAKE CIRCULATION

RICHARD H. GALLAGHER* and STEVENS T. K. CHAN†
Cornell University, Ithaca, N. Y., U.S.A.

(*Received* 14 *December* 1972)

Abstract—The finite element method is applied to the analysis of the wind-driven circulation of variable-depth, shallow, homogeneous lakes. Attention is concentrated upon higher-order descriptions of the flow phenomena within the individual elements and upon the use of these higher order functions in the definition of curved element boundaries (isoparametric elements.) Numerical results are presented for a rectangular basin, for which alternative results are available from both first-order finite element representations and finite difference analyses, and also for Lake Ontario, for which only the first-order finite element solution is available for comparison. These comparisons confirm the accuracy and efficiency of the finite element method in this field of application.

1. INTRODUCTION

The finite element method has drawn increasing attention as a numerical analysis tool for problems in fluid dynamics. The reasons for this growth of interest include the following: (1) irregular boundaries can be treated accurately without computational difficulties or changes in formulation of the method or computer program, (2) wide use can be made of universally-available general-purpose programs[1] which are virtually unlimited in the size of problem they can handle, and (3) inhomogeneous or variable properties of the problem can be easily taken into account.

Because of these advantages the finite element method is especially attractive as a method of analysis of lake circulation problems. Natural lakes of course feature irregular boundaries and the phenomenon being described is so complex in form that any numerical analysis procedure will necessarily entail hundreds, or perhaps thousands, of unknowns. Proper description of thermal stratification introduces the need to deal with spatially-varying physical properties.

The authors have presented, in ref. 2, a finite element formulation and numerical results for the wind-driven circulation of variable-depth, shallow, homogeneous lakes. The governing differential equation of this problem was taken to be that derived by Liggett and Hadjitheodourou (ref. 3). The independent variable in this equation is a specially defined stream function which we will refer to as the "flow parameter". Using the method of weighted residuals (ref. 4), with the Galerkin criterion in the selection of the weighting function, the integral form necessary for the construction of a finite element representation was constructed. To this point the work described in ref. 2 is quite general as a basis for finite element analysis of the subject problem. The adopted representation of the element flow parameter was then limited, for the purpose of generating numerical results, to a first order (linear) function on a triangular domain.

* Professor of Civil Engineering.
† Research Assoc[illegible] School of Civil and Environmental Engineering.

The present paper extends the work of ref. 2 to higher-order finite element repr tions, where the term "higher-order" refers to the level of sophistication in b geometric description of the element and the assumed representation of the flow par Evidence from the extensive finite element analysis experience in structural mechar shown that higher-order representations are in general more efficient than the simplest of element formulation. There are, naturally, limits on the degree of sophistica higher order formulations. A more significant motivation for the use of higher order el arises in the lake circulation problem. Since no experimental evidence exists for driven circulation of large lakes (we refer in this to uncertainty with respect to both th input and the lake response) it is essential that some economically feasible means checking be available in any numerical analysis technique. Such checks are made av in finite element analysis through comparison of alternative solutions obtained with d types of elements.

Two distinct classes of higher-order element representation are treated in this pap first is a direct expansion of the degree of polynomial representation of the flow par within a triangular element domain. The simplest triangular element employs a linea the element considered here describes the flow parameter with use of a cubic polyr The second type of element is intrinsically rectangular and employs quadratic re tation of the flow parameter. This same field is also employed to describe the boun the element in a special curvilinear coordinate system. This is a particular case of t parametric element approach to the formulation of elements with curved boun wherein the same ("iso") parameters are used to describe both the behavior of the e and its geometry.

The paper is organized as follows. First, a brief description is given of the formula the subject lake circulation problem as a finite element analysis problem. A more d development of this work as well as a review of efforts on finite element lake circ analysis by other investigators, can be found in ref. 2. Then, the formulation of the tw of higher order elements is detailed and attention is given to certain questions the regarding boundary conditions. Finally, numerical results are presented for recta basin problem, defined first by Liggett and Hadjitheodourou[3] and Lake Ontari importance of the rectangular basin analysis is that comparison results are available f only the first- and higher-order finite element representations, but also from a pr finite difference solution. The significance of the Lake Ontario analysis is that verifi of the integrity of the numerical finite element solution can derive only from th contained comparison of finite element representations of different degrees of refin

2. BASIC THEORY OF FINITE ELEMENT REPRESENTATION OF SUBJECT PROBLEM

Detailed development of the governing differential equations of the subject probl formulated by Liggett and Hadjitheodourou[3], and of the integral form associated v finite element representation, is beyond the scope of this paper. Rather, we define th physical problem and the final form of the governing equations.

A cross-section of the type of lake under study is pictured in Fig. 1. We fix the or coordinates at the surface of the lake with z measured upwards. In accordance w assumption of shallowness (i.e. hydrostatic pressure distribution), $D \ll L$. The eddy vi (η) and Coriolis parameter (f) are assumed constant in the form[illegible]on of the diffe relationships. The distribution of pressure is assumed to be hydr[illegible]atic and surfac

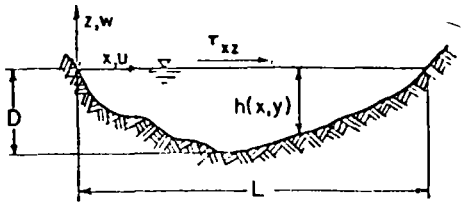


Fig. 1. Representative lake cross section

stresses (τ_{xz}, τ_{yz}) are prescribed. In order to linearize the problem the Rossby number (ratio of inertial forces to rotational forces) is taken to be small. The values of the lake depths ($h(x, y)$) are assumed to be the actual depths under the assumed wind stress or, alternately, that the equilibrium depths are a sufficient approximation to the actual depth under the assumed wind stress. The x, y-plane coincides with the water surface and $w \equiv 0$ at $z = 0$.

With these assumptions, appropriate forms of the momentum equations are constructed, the continuity equation is invoked, and after integration in the z-direction and introduction of a stream function (ψ) there is obtained

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + A\frac{\partial\psi}{\partial x} + B\frac{\partial\psi}{\partial y} + C = 0 \tag{1}$$

with the boundary condition that ψ is constant on the shoreline. The stream function is related to the average velocities ($\bar{u}$, $\bar{v}$) by

$$\bar{u} = \frac{1}{h}\frac{\partial\psi}{\partial y}, \qquad \bar{v} = -\frac{1}{h}\frac{\partial\psi}{\partial x} \tag{2}$$

The coefficients A, B, and C in the equation are functions of the planform location in x and y coordinates (more specifically, functions of the lake bottom topography) as defined in ref. 3, and C depends on the wind shear stresses as well.

The finite element representation of (1) is obtained by the method of weighted residuals. This concept assumes that an approximate representation of the independent variable, designated by $\bar{\psi}$, which does not satisfy the governing differential equation, will be chosen. In the present case this approximating *trial function* is of the form

$$\bar{\psi} = \sum_{i=1}^{n} N_i\psi_i = {}_{\llcorner}N_{\lrcorner}\{\psi\} \tag{3}$$

where ψ_i is a particular value of the independent variable and generally refers to such a value at the point i, $i = 1, \ldots n$, and the coefficients N_i, which are functions of the x and y coordinates, are termed *shape functions*.

Designating the governing differential equation (1) as $L(\psi) = 0$, we note that due to the approximate nature of $\bar{\psi}$ we have

$$L(\bar{\psi}) = R \neq 0 \tag{4}$$

where R is a residual value. Since the governing differential equation cannot be satisfied pointwise throughout the domain (V) of the problem we can seek its satisfaction in the sense of a weighted average over the domain, i.e.

$$\int_V L(\bar{\psi})\phi \, dV = 0 \tag{5}$$

where ϕ is the weighting function, which may be specified in one of any number of fi Here, we choose the Galerkin form in which the coefficients (N_i) of the trial functio employed. Each distinct trial function leads to a separate algebraic equation, as fol First, we substitute (3) into $L(\psi)$, as given by equation (1) and designate the integr defined by equation (5), where now $\phi = N_i$. We obtain, for all N_i, $i = 1, \ldots n$, in element

$$\iint_\Delta \{N\}\left[\left(\frac{\partial^2{}_{\llcorner}N_{\lrcorner}}{\partial x^2} + \frac{\partial^2{}_{\llcorner}N_{\lrcorner}}{\partial y^2} + A\frac{\partial{}_{\llcorner}N_{\lrcorner}}{\partial x} + B\frac{\partial{}_{\llcorner}N_{\lrcorner}}{\partial y}\right)\{\psi\} + C\right] dx\, dy = \{0\}$$

where $\{N\} = N^T$ is a $1 \times n$ column vector containing the shape functions N_i.

Next, one applies integration by parts in the plane, reducing the order of the deriv appearing in the integral and introducing the boundary terms into the resulting int This leads to the following system of algebraic equations

$$[k^e]\{\psi\} = \{r^e\} + \oint^e$$

where

$$[k^e] = \iint_\Delta \left(-\frac{\partial\{N\}}{\partial x}\frac{\partial{}_{\llcorner}N_{\lrcorner}}{\partial x} - \frac{\partial\{N\}}{\partial y}\frac{\partial{}_{\llcorner}N_{\lrcorner}}{\partial y} + A\{N\}\frac{\partial{}_{\llcorner}N_{\lrcorner}}{\partial x} + B\{N\}\frac{\partial{}_{\llcorner}N_{\lrcorner}}{\partial y}\right) dx\, dy$$

$$\{r^e\} = -\iint_\Delta \{N\}C\, dx\, dy$$

and $\oint^e$ symbolizes a boundary integral.

Due to the terms $A\{N\}$ ($\partial_{\llcorner}N_{\lrcorner}/\partial x$) and $B\{N\}$ ($\partial_{\llcorner}N_{\lrcorner}/\partial y$) these algebraic equation nonsymmetric. Also, as noted previously, the coefficients A and B are functions of x a

The equations of the complete lake are constructed from the equations of the elemen imposing the condition of stream function continuity at each element joint, which is syn mous with simple addition of all coefficients ($k_{ij}{}^e$ and $r_j{}^e$) with like subscripts. Thus, th set of equations is of the form

$$[K]\{\psi\} = \{R\}$$

where

$$K_{ij} = \sum k_{ij}{}^e$$

$$R_i = \sum r_i{}^e + \sum \oint^e$$

and the summations range over all elements with terms with the subscripts i and j.

After solution of equation (10) for $\{\psi\}$, other variables, such as averaged velocities, sure gradients, and velocities at different depth levels can be subsequently evaluate back substitutions into equations presented in ref. 3.

We next examine, in the following two sections, the choice of shape functions N_i (equ 3) for two classes of higher order elements and the use of these functions in the develop of element equations from equations 8 and 9.

3. HIGHER ORDER TRIANGLE

The triangle holds a special place in finite element analysis due to its association "complete" polynomials of lower order. This point is illustrated by the array of the c cients of the polynomial series in the form of a Pascal triangle (Fig. 2). A complete l

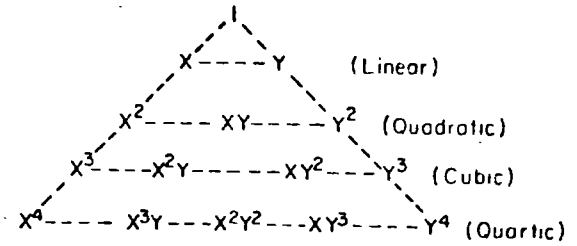


Fig. 2. Pascal triangle and relationship to higher-order elements

function is identified with the simplest triangle (the location of the coefficients in the Pascal triangle identifies joints in the element), a quadratic function corresponds to a triangle with joints at both the vertices and the midpoints of the sides. In this way we can perceive a "family" of triangles with no upper limit on the degree of polynomial employed. The functions of degree higher than one are collectively termed "higher-order" functions and when an element formulation is based on such functions it is called a "higher-order" element.

The advantage of a higher order element derives from two principal considerations. It is possible to write the higher-order function directly in "shape function" form. The Pascal triangle identifies a polynomial series representation of ψ, where the coefficients of the series do not have the physical significance of shape functions. For more general shapes of element the transformation of the polynomial coefficients into shape function form may be an expensive operation. Secondly, explicit formulas are available for the integrals of the shape functions over the triangle domain. Both of these considerations, which are detailed in ref. 6, stem from the existence of triangular coordinates, defined as

$$(L_i)^p = \frac{(\text{Area})_i}{(\text{Area})} \tag{13}$$

where (Area) is the total area of the triangle, $(\text{Area})_i$ is the area of the triangular subregion $i (i = 1, 2, 3)$ (see Fig. 3a) and is a linear function of the x, y coordinates of point p. It

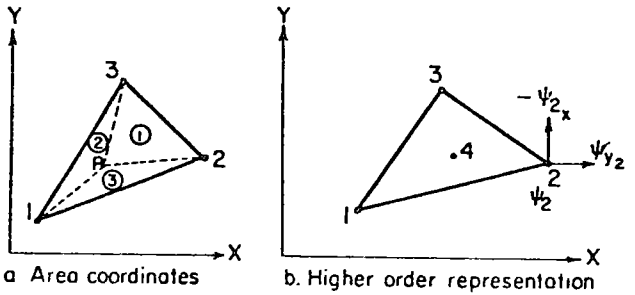


Fig. 3. Triangular elements

follows that one can construct a linear representation of a quantity such as B, the coefficient in equation (1), as follows

$$B = L_1 B_1 + L_2 B_2 + L_3 B_3 \tag{1}$$

where B_i, $i = 1, 2, 3$ are the values of B at the element vertices.

As Fig. 2 discloses, the basic higher-order triangle possesses 6 joints and the shape tion corresponds to a complete quadratic polynomial. Chan and Larock[7] have utiliz element of this form in the analysis of potential flow problems The results obtain application of the element disclosed significant improvements over those obtained linear-field elements.

We go one step further in the present approach and adopt a shape function based u complete cubic polynomial. Rather than defining element joints at the ten location scribed in Fig. 2, however, we deal with only the three vertices and a point at the centr the element (see Fig. 3.b). To accommodate the ten terms of a complete cubic we tre derivatives of ψ at the vertices ($\partial\psi/\partial x$, $\partial\psi/\partial y$, at points 1, 2 and 3) as solution param Thus, at a typical vertex i we have as parameters ψ_i, $\psi_{x_i} = (\partial\psi/\partial x)|_i$, $\psi_{y_i} = (\partial\psi/\partial y)|_i$ only solution parameter at point 4 is ψ_4.

By introducing the derivatives of ψ as solution parameters and thereby eliminating along the side of the element we reduce the bandwidth of the algebraic equations solved (equation 10). Furthermore, from equation (2), the solution parameters are di proportional to quantities of interest, $\bar{u}$ and $\bar{v}$, and so the latter are in effect directly pr out after computation of the solution.

The present triangle (Fig. 3b) has been used extensively in finite element analysis number of authors, including flow analysis problems[8]. The shape function descripti ψ, given by Felippa[9] is

$$\psi = N_1\psi_1 + N_2\psi_{x_1} + N_3\psi_{y_1} + N_4\psi_2 + \cdots N_9\psi_{y_3} + N_{10}\psi_4$$

where

$$N_1 = L_1^2(L_1 + 3L_2 + 3L_3) - 7\gamma \qquad N_6 = L_2^2(y_{12}L_1 - y_{23}L_3) + (y_{23} - y_1$$

$$N_2 = L_1^2(x_{12}L_2 - x_{31}L_3) + (x_{31} - x_{12})\gamma \qquad N_7 = L_3^2(L_3 + 3L_1 + 3L_2) - 7\gamma$$

$$N_3 = L_1^2(y_{31}L_3 - y_{12}L_2) + (y_{12} - y_{31})\gamma \qquad N_8 = L_3^2(x_{31}L_1 - x_{23}L_2) + (x_{23} - x$$

$$N_4 = L_2^2(L_2 + 3L_3 + 3L_1) - 7\gamma \qquad N_9 = L_3^2(y_{23}L_2 - y_{31}L_1) + (y_{31} - y_2$$

$$N_5 = L_2^2(x_{23}L_3 - x_{12}L_1) + (x_{12} - x_{23})\gamma \qquad N_{10} = 27\gamma$$

$$x_{ij} = x_i - x_j \quad , \quad y_{ij} = y_i - y_j \qquad \gamma = L_1L_2L_3$$

In the evaluation of the element coefficients, through integration of equations (8) an the terms A, B, and C are assumed to have linear variation over the region of the eleme given by equations of the form of (14). This is convenient since B_1, B_2, etc., the valu these quantities at the "joints", are the conventional input parameters in an analysis

One inconvenient aspect of the element formulation resulting from the use of equ (16) is the presence of point 4 in the interior of the element. This solution parameter doe join to any other so it can be eliminated from the element equation before assembly c latter into the system representation (equation 10). Thus, the element equations a order 9 × 9 immediately before assembly.

As mentioned in Section 2, consideration of boundary conditions requires that the st function is a constant, which for convenience is chosen to be zero, all along the boun In the present representation the stream function varies cubically between two adj joints and requires specification of four parameters for unique defini n along such an Thus, in addition to the two stream function values at the joints th lues of the tang derivative of the stream function at both end joints should also be zero. As in finite el

analysis by using triangular elements, the boundary is usually approximated by broken lines and hence the tangential direction of a joint is rather ambiguously defined. Hence, some further approximation is needed to choose its direction. For instance, one may consider the average direction of the two lines meeting at the same joint as its tangential direction, or alternatively, take either one indicated by the two lines as the required direction (but consistently for all the joints). The latter is adopted in the present study for the sake of simplicity. That is, the line connecting a joint and its adjacent joint in front, taken in counterclockwise order, is considered to be the tangential direction at that joint. Once the tangential direction at each boundary joint has been established, a matrix transformation is necessary in order to introduce the tangential and normal derivatives of the stream function, in place of the x- and y-derivatives, as solution parameters. This transformation can be performed either at the element level or at the system matrix level. After that, the stream function value and its tangential derivative at each of the boundary joint are set equal to zero and finally the system of equations is solved.

D. ISOPARAMETRIC ELEMENT

The four-sided isoparametric element is shown in Fig. 4. The simplest order of shape function which will describe curved boundaries, a quadratic, is chosen here. A curvilinear

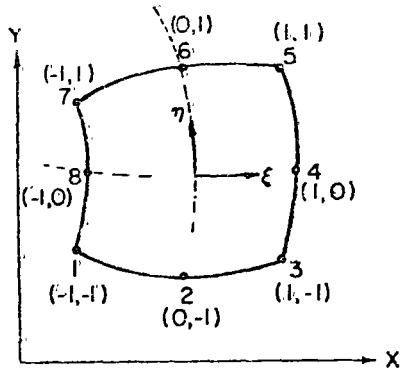


Fig. 4. Quadratic isoparametric element

coordinate system (ξ, η) is defined within the element in such a way that the corners of the element have coordinates $+1$ or -1 as indicated in the figure. Rectangular coordinates (x, y), in terms of which the location of node points are initially defined, are also established.

The stream function for the velocity is described by the following

$$\psi = N_1\psi_1 + \cdots N_i\psi_i + \cdots N_8\psi_8 \tag{17}$$

We note that in this case only the values of ψ at the joints, and not the derivatives of ψ, are chosen as unknowns. The shape functions N_i are given in ref. (10) by

$$N_i = \tfrac{1}{4}[(1 + \xi\xi_i)(1 + \eta\eta_i) - (1 - \xi^2)(1 + \eta\eta_i) - (1 + \xi\xi_i)(1 - \eta^2)] \quad \text{for } i = 1, 3, 5, 7$$

$$N_i = \tfrac{1}{2}(1 - \xi^2)(1 + \eta\eta_i) \quad \text{for } i = 2, 6 \tag{18}$$

The validity of these expressions is confirmed in two ways. First, the N_i's must take values such that $\psi = \psi_j$ when ψ is evaluated for the coordinates of point j (i.e., $N_i = 1$, $i = j$, $N_j = 0$ for $i \neq j$). Secondly, when evaluating ψ along an edge it should be found to b function of only the values of ψ_i along that same edge. This means that ψ will be continu across element boundaries and in forming a geometric idealization with curved bounda the geometry will be continuous if the same shape functions are adopted, i.e.

$$x = N_1 x_1 + \cdots N_i x_i + \cdots N_8 x_8$$

$$y = N_1 y_1 + \cdots N_i y_i + \cdots N_8 y_8$$

Now, it is recalled from equation (8) that evaluation of the element coefficients invo the derivatives of the shape functions with respect to x and y and integration over the of the element. The above shape functions are defined in terms of ξ and η, however, so the transformation to x and y must be established. From the chain rule of differentiat for a typical shape function N_i

$$\begin{Bmatrix} \dfrac{\partial N_i}{\partial \xi} \\ \dfrac{\partial N_i}{\partial \eta} \end{Bmatrix} = \begin{bmatrix} \dfrac{\partial x}{\partial \xi} & \dfrac{\partial y}{\partial \xi} \\ \dfrac{\partial x}{\partial \eta} & \dfrac{\partial y}{\partial \eta} \end{bmatrix} \begin{Bmatrix} \dfrac{\partial N_i}{\partial x} \\ \dfrac{\partial N_i}{\partial y} \end{Bmatrix} = [J] \begin{Bmatrix} \dfrac{\partial N_i}{\partial x} \\ \dfrac{\partial N_i}{\partial y} \end{Bmatrix}$$

where $[J]$, the *Jacobian* matrix, is evaluated by differentiation of (19). Also

$$dA = |J|\, d\xi\, d\eta$$

where $|J|$ symbolizes the determinant of $[J]$.

Equations (20) and (21) give the necessary basis for evaluation of (8). The resulting int is too complicated for explicit evaluation so that numerical integration must be emplo Gaussian integration was applied in the present case, with 3 Gaussian points in each direct The coefficients A, B and C were evaluated at the origin of ξ-η coordinates and were assu to be constant throughout the element. It should be noted that in the presence of det data for a given physical problem it would be feasible to evaluate these quantities at eac the numerical integration points, and thereby obtain a better approximation to the equa coefficients. However, computation time is expected to increase considerably as the exp sions for these coefficients are rather involved.

E. NUMERICAL RESULTS

To study the feasibility of the present approach, two problems, presented in ref. 2, again analyzed by using the higher-order elements described herein. For both probl fewer isoparametric elements, and henceforth a reduction in the number of solution p meters, have been used to demonstrate the merits of high-order element over the si linear stream function representation. The same idea could have been applied to the f field representation but it was decided to use the same element gridworks adopted in linear field so as to gain some idea about how solution eventually converges with th crease in number of solution parameters.

The first problem, shown in Fig. 5, is an idealized rectangular basin oriented in a n south direction with a length four times its width. Wind was assumed to blow from Sou North. Values employed in the actual computations are shown in Fig. 5.

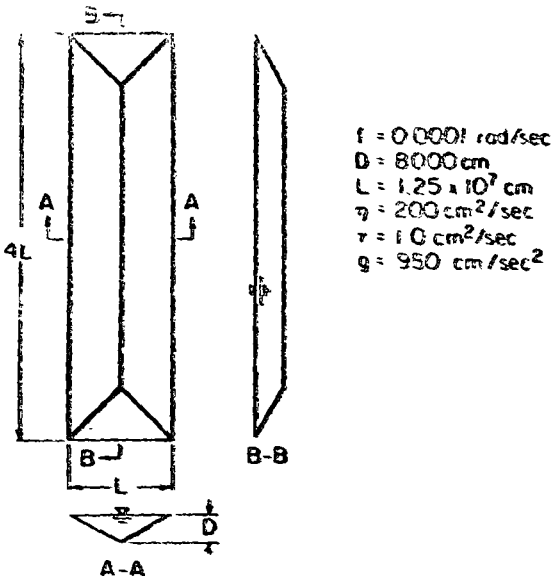


Fig. 5. Rectangular lake

As mentioned in ref. 3, a zero depth represents a computational singularity. Hence, only the flow region bounded by a contour of 5 per cent maximum depth, as used in ref. 2, is considered and the effects from flow exterior to the boundary are assumed to be negligibly small. This choice of the flow region, rather than using the actual shoreline boundary with a non-zero depth assigned along it, is preferred because the bottom topography of the lake is retained. It is to be noted that although the lake is geometrically symmetric about x- and y-axes, this property of symmetry does not apply to the circulation behavior being calculated. Therefore the entire lake must be considered. Fig. 6 shows the finite element gridworks for a quadrant of the lake. For the entire lake, 360 triangular elements (either linear or cubic) with 209 joints, or 40 isoparametric elements (quadratic) with 149 joints, have been used in the numerical computation.

Figures 7 and 8 show results for the stream function ψ and the magnitude of surface velocity at selected sections, as predicted by various finite element representations. Also shown are the finite difference results from ref. 3, where 1701 equally-spaced pivotal points were used. For the finite element representations, 209 solution parameters were used for the linear field, 149 for the quadratic field, and 627 for the cubic field approximation. The finite element and finite difference results are seen to be in close agreement. Also, it is noted that results obtained by using isoparametric elements compare well with others, in spite of the fact that only an amount of 72 per cent in solution parameters, compared to the linear field representation, has been used. This fact seems to confirm that higher-order element representation is more desirable, regarding accuracy and efficiency, over the simple linear field representation. Of course, our results could have been improved further if the coefficients A, B, and C in the governing equation had been evaluated at all Gaussian points instead of only one point for each element as presently carried out. The results from cubic field representation, as expected, are far more accurate and coincide almost exactly with the finite difference results based on a much finer gridwork.

The seco roblem considered is the prediction of circulation of Lake Ontario oue to a

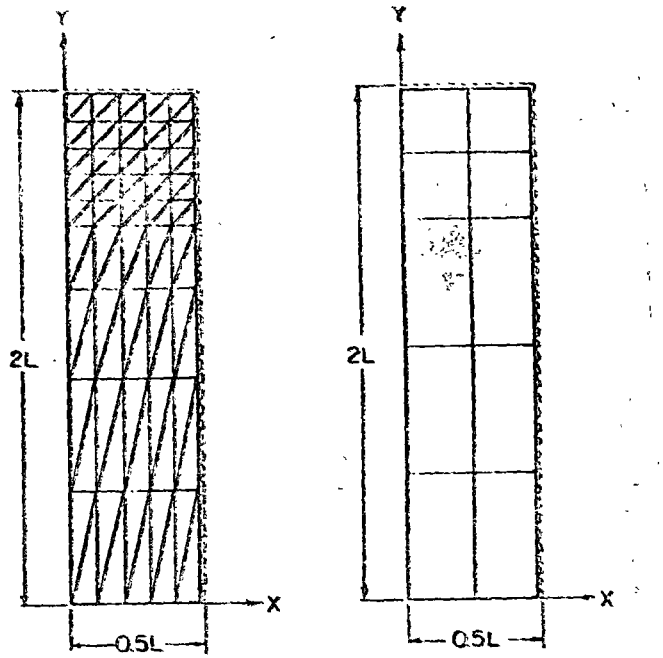


Fig. 6. Finite element representation of rectangular lake

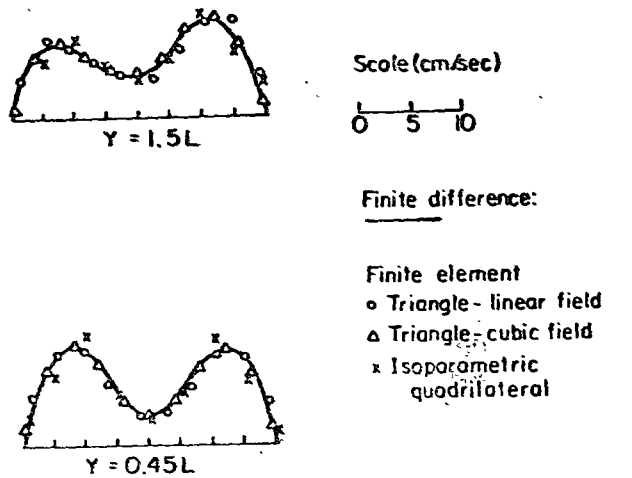


Fig. 7. Comparison of total velocity solutions on rectangular lake surface at representative sections

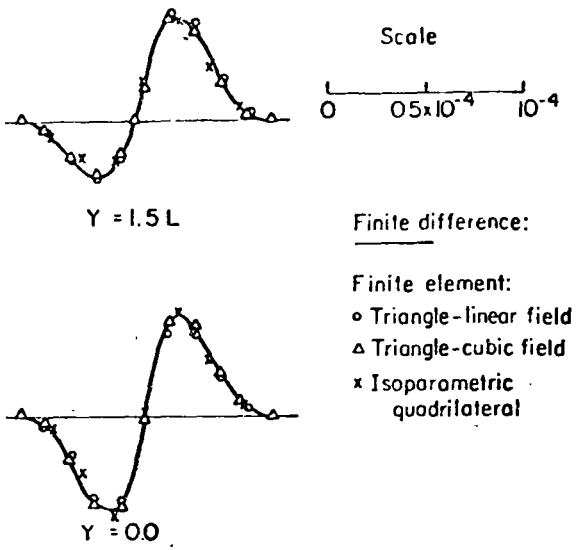


Fig. 8. Comparison of stream function solutions for rectangular lake at representative sections

wind shear prevailing in the NE direction with 7° to the East (the local average direction at Rochester in February). The same numerical values of f, η, τ, g as used in the previous example were used in the present case. The characteristic length (L), which can be chosen arbitrarily, was set to be 2×10^7 cm, while the characteristic depth (D) was chosen to represent the maximum nodal depth after the lake had been discretized with finite elements. The value of D may vary slightly with the gridwork. For instance, it is 22500 cm for the triangular element gridwork and 21650 cm for the isoparametric representation. Figures 9 and 10 show the gridworks employed in the present analysis: 561 triangular elements with 323 joints, or 70 isoparametric elements with 257 joints. The geometry and bottom topography of the lake described in ref. 5 was used herein to define the finite element representations. Again, to avoid computational singularity, the flow region studied is the one bounded by the contour line having depth of water of 12 m, not the actual shoreline.

Figures 11 and 12 show the stream function values and the magnitude of surface velocity

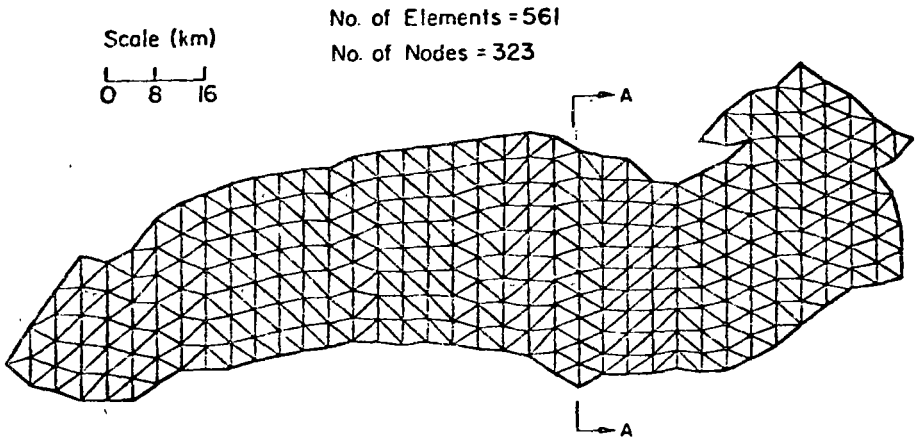


Fig. 9. Finite element representation of Lake Ontario by triangular elements

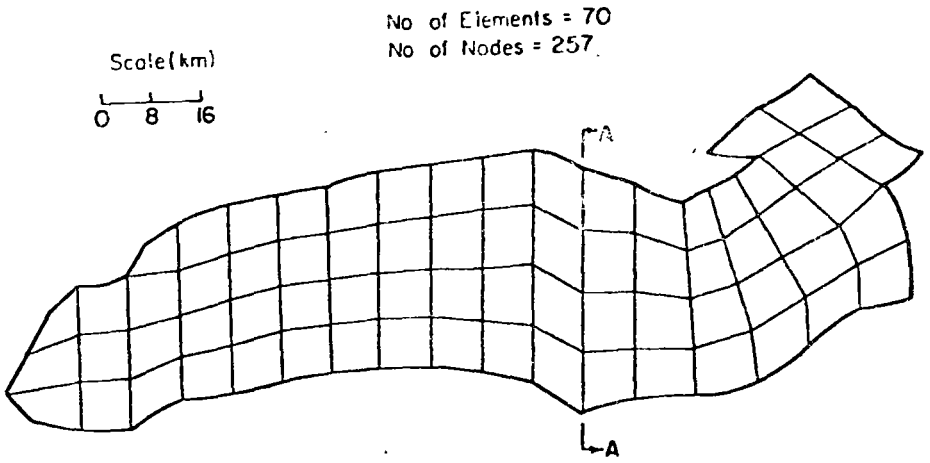


Fig. 10. Finite element representation of Lake Ontario by isoparametric elements

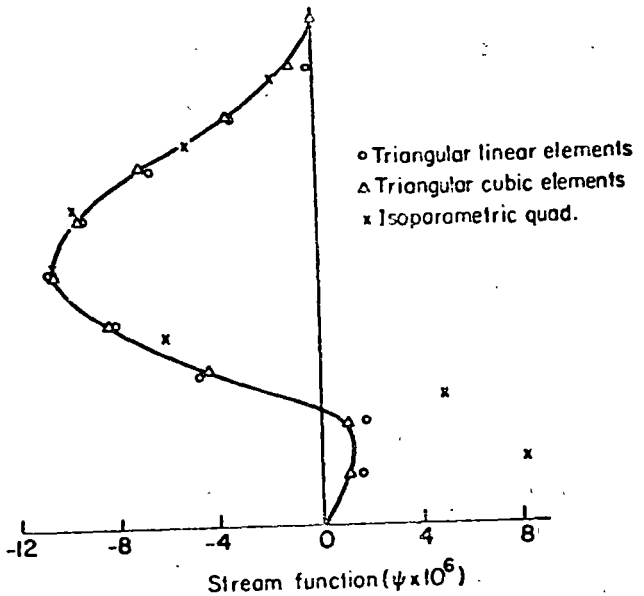


Fig. 11. Comparison of stream function solution for Lake Ontario at section *A-A*

at a representative section of the lake. These results are seen to be in good agreemer general. Near the south shoreline, however, some discrepancies in the stream function e between the isoparametric element prediction and those obtained by the other two repres ations. The results obtained by the former are believed to be inaccurate because a ra coarse gridwork had been used, and as a consequence, the bottom topography might have been properly represented in that region. This inaccuracy, however, does not see affect seriously the surface velocity prediction in that region, as seen in Fig. 12. This suggests that surface velocity distribution, which is mainly affected by the surface and rounding conditions, is less influenced by the bottom topography than the stream funct

No comparison results by other methods are available for this problem. Although

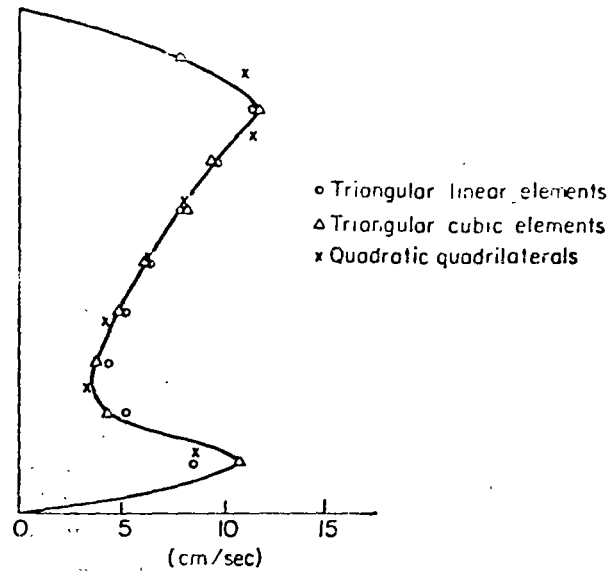


Fig. 12. Comparison of total velocity solutions on Lake Ontario surface at section *A–A*

correct Coriolis parameter was used, no serious attempt was made to choose a physically accurate eddy viscosity or to account for ice formation or variation of wind stress. It is unlikely that field measurements of the form necessary for comparison purposes will be available in the foreseeable future. Large-scale modeling is a promising alternative source of comparison data but no such data yet exists for this lake and when they are obtained it is to be expected that limitations on representation of the pertinent dimensionless ratios (ref. 11) will require somewhat different conditions on the comparison analysis than those employed herein.

F. CONCLUDING REMARKS

The results presented in this paper demonstrate that higher-order and isoparametric elements represent important components of the most effective utilization of the finite element method in lake circulation analysis. The improved accuracy of the triangle with cubic stream-function field is gained without significant increase in formulative cost over that involved in the simplest triangle. Additionally, the use of derivatives of the stream function as solution parameters is a convenience in the treatment of boundary conditions involving prescribed inflow or outflow rate. It is questionable, however, that still higher-order representations (e.g., based on quintic functions) would be more efficient because for these the element formulative cost is no longer insignificant. Also, the large number and type of solution parameters per element may prove awkward. In any case, realistic assessments of "optimal" degree of higher-order representation are needed.

Isoparametric element representation would appear to be of special importance in the finite element analysis of natural lakes, which possess irregular shorelines. The representation of such lakes can be accomplished wit'. large numbers of straight-sided elements but in that case many elements and solution parameters are employed solely for geometr' representation. Th oparametric concept enables each element to serve the functions o. both geometric and behavior representation. This factor will be of overriding importance as

the theoretical representation is extended to three dimensions and more complex phen e.g. thermal-fluid interaction. It should be noted that although the four-sided ele popular in isoparametric representation, it is also possible to cast the triangle in th Clearly, the question of "optimal" degree of shape function is once again imp Evidence[12] tends towards the desirability of second- or third-degree functions.

Acknowledgment—Work described in this paper was supported by the National Science Foundati Research Grant GK-23992.

REFERENCES

1. Gallagher, R. H., "*Large-Scale Computer Programs for Structural Analysis*" *in General Purp Element Computer Programs*, P. V. Marcal, Editor. ASME Special Publication, November 19
2. Gallagher, R. H , Liggett, J. and Chan, S., "Finite Element Analysis of Circulation in Variab Shallow Homogeneous Lakes", *J. Hyd. Div.*, ASCE, 99, 1973.
3. Liggett, J. A., and Hadjitheodorou, C., "Circulation in Shallow, Homogeneous Lakes", Proc *J. Hydraulics Div.*, 95, No. HY2, 1969, 609–620.
4. Finlayson, B. A., "*The Method of Weighted Residuals and Variational Principles*", Acaden N.Y., 1972.
5. Boyce, F. M., "Lake Ontario Digitized Bathymetry", 2nd revised edition, Canada Center fc Waters (private communication).
6. Silvester, P., "High-Order Polynomial Triangular Elements for Potential Problems", *Int. Sci.* 7, 849–861, 1969.
7. Chan, S T. K. and Larock, B. E., "Potential Flow Around a Cylinder Between Parallel Walls Element Methods", *J. Eng. Mech Div*, ASCE, 98, No. EM5, 1317–1322, 1972.
8. Argyris, J. H., Mareczek, C. and Scharpf, D., "Two- and Three-Dimensional Flow Using F ments", *Aeronaut J.*, 73, 1969, 961–964.
9. Felippa, C. A., "Refined Finite Element Analysis of Linear and Nonlinear Two-Dimensional Str Report No. SESM 66–22, Structural Engrg Lab., Univ. of Calif. at Berkeley, 1966.
10. Ergatoudis, I., Irons, B and Zienkiewicz, O., "Curved Isoparametric Quadrilateral Elements f Element Analysis", *Int. J. Solids Struct.*, 4, 1968, 31–42.
11. Rumer, R. and Hoopes, J. A., "Modelling Great Lakes Circulations", Report, Wiscons. Resources Center, 1971.
12. Zienkiewicz, O. C., "*Isoparametric and Allied Numerically Integrated Elements—A Review*" ONR International Symposium on Numerical and Computer Methods in Structural Mech. Robinson, Ed., Academic Press, New York, 1972.

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

USOS DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPRESION EN AGUAS, COSTERAS LAGOS Y RIOS

METODOS DE LOS RESIDUOS PESADOS

MAYO, 1978.

APPLIED MECHANICS REVIEWS

VOL. 19, NO. 9 SEPTEMBER 1966

Gustavo Argaiz M. Mexico 1971

The Method of Weighted Residuals—A Review

B. A. FINLAYSON* AND L. E. SCRIVEN

UNIVERSITY OF MINNESOTA,
MINNEAPOLIS, MINNESOTA

Abstract

The method of weighted residuals unifies many approximate methods of solution of differential equations that are being used currently. This review presents the basic method in its historical context and shows some of the many possible modifications that have been used throughout the past fifty years. The relationship between the Galerkin method, which is one version of the method of weighted residuals, and variational methods is outlined. Also included is an extensive listing of published applications of the method of weighted residuals.

Introduction

The method of weighted residuals is an engineer's tool for finding approximate solutions to the equations of change of distributed systems. Experience and intuition can be distilled into a reasonable and sometimes quite accurate first guess, from which it is possible to proceed to successively improved approximations. The analytical form of the approximate solution is often more useful than solutions generated by numerical integration, and the approximate solution usually requires less computation time to generate. The method is applicable to nonlinear and non-self-adjoint problems—one of its most attractive features.

The method of weighted residuals (MWR) includes many approximation methods that are being used currently. It provides a vantage point from which it is easy to see the unity of these methods as well as the relationships between them. This review, after outlining application of the basic method to initial-value, boundary-value, and eigenvalue problems, surveys the history of major contributions to the subject and discusses some of the many modifications of the basic method. The review concludes with a listing of applications of weighted residual methods to problems arising in applied mechanics and related fields. Four practical aspects of MWR in need of further research are identified.

*Present address: Office of Naval Research, Washington, D. C.

A. Basic Method

The best available treatments of MWR have been those by Crandall [1], who coined the name method of weighted residuals, Ames [2], and Collatz [3], who calls these methods error-distribution principles. The following outline parallels their treatments, in places contrasting them and elaborating on them.

Given a system of differential or integro-differential equations of change and constitutive relations, boundary conditions representing the interactions between the system and its surroundings, and initial conditions representing some base state of interest, the general approach is to assume a trial solution whose functional dependence on position is chosen, but which includes undetermined functions of time. The latter are found by requiring that the trial solution satisfy the differential equation in some specified approximate sense.

Initial Value Problem

Consider the differential equation for $(u(\mathbf{x}, t)$:

$$N(u) - \frac{\partial u}{\partial t} = 0 \quad \mathbf{x} \text{ in } V,\ t > 0 \tag{1}$$

where $N(\cdot)$ denotes a general differential operator involving spatial derivatives of u, V is a three-dimensional domain with boundary S, and t represents time. Suppose the initial and boundary conditions are

$$\begin{aligned} u(\mathbf{x}, 0) &= u_0(\mathbf{x}), \quad \mathbf{x} \text{ in } V \\ u(\mathbf{x}, t) &= f_S(\mathbf{x}, t), \quad \mathbf{x} \text{ on } S \end{aligned} \tag{2}$$

Assume a trial solution of the form

$$u^*(\mathbf{x}, t) = u_S(\mathbf{x}, t) + \sum_{i=1}^{N} c_i(t) u_i(\mathbf{x}, t) \tag{3}$$

where the approximating functions, u_i, are prescribed and satisfy the boundary conditions

$$u_S = f_S, \quad u_i = 0, \quad \mathbf{x} \text{ on } S \tag{4}$$

Then u^* satisfies the boundary conditions for all functions $c_i(t)$. It is not necessary that the trial solution

be linear in the c_i, but such a choice is usually made for simplicity; no systematic study of alternatives has been reported, so far as the authors know. The differential equation residual and initial residual,

$$R(u^*) \equiv N(u^*) - \frac{\partial u^*}{\partial t} \tag{5}$$

$$R_0(u^*) \equiv u_0(\mathbf{x}) - u_s(\mathbf{x}, 0) - \sum_{i=1}^{N} c_i(0)\, u_i(\mathbf{x}, 0) \tag{6}$$

are measures of the extent to which the function u^* satisfies the differential equation and initial conditions, respectively. As the number N of approximating functions u_i is increased in successive approximations, one hopes the residuals will become smaller; the exact solution is obtained when both residuals are identically zero. As an approximation to this ideal, the weighted integrals of the residuals are set equal to zero:

$$\begin{aligned} \langle w_j;\, R(u^*) \rangle &= 0 \\ \langle w_j;\, R_0(u^*) \rangle &= 0 \end{aligned} \qquad j = 1, 2, \ldots, N \tag{7}$$

where

$$\langle w, v \rangle \equiv \int_v wv\, dV \tag{8}$$

represents a spatial average or inner product and w_j is a prescribed weighting function. If u^* is the exact solution, Equations (7) are satisfied regardless of the choice of weighting functions.

The weighting functions can be chosen in several different ways, and each choice corresponds to a different criterion in MWR. Once the choice is made, Equations (7) become a set of N first-order ordinary differential equations in the N unknowns $c_i(t)$. For the linear problem

$$\frac{\partial u}{\partial t} = L(u) \tag{9}$$

with approximating functions u_i and u_s that do not themselves depend on time, Equations (7) become simply

$$\sum_{i=1}^{N} \frac{dc_i}{dt} \langle w_j;\, u_i \rangle = \sum_{i=1}^{N} c_i \langle w_j;\, L(u_i) \rangle + \langle w_j;\, L(u_s) \rangle \tag{10}$$

or, in matrix notation

$$\overline{\overline{A}}\, \frac{d\bar{c}}{dt} = \overline{\overline{B}}\, \bar{c} + \bar{b}. \tag{11}$$

The solution to these equations is substituted into Equation (3) to give the approximate solution to the problem. Successive approximations are obtained by increasing N and solving Equation (10) anew. The convergence of successive approximations gives a clue, but not necessarily a definitive one, to the reasonableness of the approximation.

Boundary Value and Eigenvalue Problems

The method is equally applicable to steady-state and eigenvalue problems. For steady-state problems, the c_i are constants rather than functions of time; for linear problems they are determined as solutions to

$$\overline{\overline{B}}\, \bar{c} = -\bar{b}. \tag{12}$$

For nonlinear boundary-value problems it may be useful to assume trial solutions of a more general form than Equation (3), viz.:

$$u^*(\mathbf{x}) = \phi\left(\{c_i\}, u_i(\mathbf{x})\right) \tag{13}$$

For the linear eigenvalue problem

$$L(u) - \lambda u = 0 \tag{14}$$

the approximate solution is determined by

$$\sum_{i=1}^{N} c_i \left\{ \langle w_j;\, L(u_i) \rangle - \lambda \langle w_j;\, u_i \rangle \right\} \equiv \sum_{i=1}^{N} c_i (A_{ji} - \lambda B_{ji}) = 0 \tag{15}$$

and this set of equations has a non-trivial solution only if

$$\det (A_{ji} - \lambda B_{ji}) = 0. \tag{16}$$

The values of λ for which this is true are the approximations to the first N eigenvalues λ_k.

Weighting Functions

The choice of the weighting functions, w_j in (7), corresponds to various criteria in MWR: the historical relationship of the criteria is portrayed in Table I.

In the collocation method, due to Frazer, Jones, and Skan [4], the weighting functions are the Dirac delta functions

$$w_j = \delta(\mathbf{x}_j - \mathbf{x}); \tag{17}$$

TABLE I
HISTORY OF APPROXIMATE METHODS

Date	Investigator	Method
1915	Galerkin [10]	Galerkin method
1921	Pohlhausen [18]	Integral method
1923	Biezeno and Koch [5]	Subdomain method
1928	Picone [9]	Method of least squares
1932	Kravchuk [17]	Method of moments
1933	Kantorovich [30]	Method of reduction to ordinary differential equations
1937	Frazer, Jones, and Skan [4]	Collocation method
1938	Poritsky [31]	Method of reduction to ordinary differential equations
1940	Repman [55]	Convergence of Galerkin's method
1941	Bickley [12]	Collocation, Galerkin, least squares for initial-value problems
1942	Keldysh [57]	Convergence of Galerkin's method, steady-state
1947	Yamada [16]	Method of moments
1949	Faedo [59]	Convergence of Galerkin's method, unsteady-state
1953	Green [60]	
1956	Crandall [1]	Unification as method of weighted residuals

the differential equation is then satisfied exactly at the N collocation points, x_j. As N is increased, the residual vanishes at more and more points and presumably approaches zero throughout V.

If the weighting functions are

$$w_j = \begin{cases} 1 & \mathbf{x} \text{ in } V_j \\ 0 & \mathbf{x} \text{ not in } V_j \end{cases} \tag{18}$$

then the differential equation is satisfied on the average in each of the N subdomains, V_j; this is the subdomain method [5, 6]. If the V_j are disjoint (which they need not be), the size of one or more subdomains decreases as N is increased, with the result that the differential equation is satisfied on the average in smaller and smaller regions, and presumably the residual approaches zero everywhere. It was Biezeno's presentation [7] of the subdomain method at the First International Congress of Applied Mechanics which prompted Courant's remark [8] that led Crandall to choose the name, "method of weighted residuals." The authors' translation of Courant's remark reads:

> "Mr. Courant (Göttingen) indicated afterward that the method advanced by Mr. Biezeno can be viewed from the standpoint of the calculus of variations in the following manner. If a differential equation, as it arises for example in a variational problem, must be satisfied, then we can express it so that the left side of the differential equation, multiplied with an arbitrary function and then integrated, must give us the value zero (vanishing of the first variation). Instead of taking an arbitrary function, we can also take infinitely many determined functions, if these only form a so-called complete function system for the region in question. The piecewise constant functions advanced by Mr. Biezeno are indeed just an especially simple special case of such a complete function system."

The least-squares method, which seems to have been first presented for this type of application by Picone in 1928 [9], uses the weighting functions $\partial R(u^*)/\partial c_j$. The corresponding interpretation is that the mean square residual

$$I \equiv \int_V [R(u^*)]^2 \, dV \tag{19}$$

is minimized with respect to the constants c_j.

In the Galerkin method [10], developed in 1915 as the first criterion of what is now known as the method of weighted residuals,* the weighting functions w_j are just the approximating functions of u_j. The approximating functions are often members of a complete system of functions, although this property, required for mathematical purposes, is sometimes ignored in practice. The Galerkin method then can be interpreted as making the residual orthogonal to members of the complete set.

A fundamental property (sometimes the definition) of a complete system of functions is that a piecewise continuous function can be orthogonal to each and every member only if the function is identically zero. In the approximation scheme outlined above, the residual is usually continuous (depending upon the differential operator and the choice of approximating functions), and hence the residual can vanish only if it is orthogonal to each member of a complete system of functions. Of course in practice the residual is made orthogonal to no more than a modest, finite number of the members of a complete set. In the original Galerkin method, developed in the study of elastic equilibrium and stability of rods and plates, Galerkin used trial solutions with unknown constant coefficients. Now many similar techniques are often referred to as the Galerkin or generalized Galerkin method: (i) the one given above in which $c_j = c_j(t)$ for time-dependent problems [12, 1]; (ii) one in which trial solutions are of the more general form $u^* = f(\mathbf{x}, \{c_j\})$ with weighting functions $\partial f/\partial c_j$ [13]; and (iii) one in which weighting functions are of the form $K(u_j)$, rather than u_j, where K is a specified differential operator [14, 15].

The method of moments is similar to the Galerkin method except that the residual is made orthogonal to members of a system of functions which need not be the same as the approximating functions. Both methods are combined under the single name of orthogonalization methods by Collatz [3]. Yamada [16] and Kravchuk [17] applied the method of moments to ordinary differential equations by using the weighting functions $\{x^n\}$ regardless of the choice of approximating functions. For the first approximation, the weighting function is unity, and the method of moments in this case is equivalent to the subdomain method and is usually called the integral method, or von Kármán–Pohlhausen method [18, 19]. For the integral method, reviewed in detail by Goodman [20], the differential equation is satisfied on the average over the domain of interest.

Boundary Methods

In the foregoing it is presumed that the trial solution satisfies the boundary conditions but not the differential equation. The converse situation can also be treated: the differential equation is satisfied but the boundary conditions are not. Trial solutions of this sort lead to boundary methods, as they are called by Collatz [3]; the procedures are analogous to those above, but with the spatial average, Equation (8), replaced by an average over the boundary.

Mixed Methods

The intermediate situation can also be handled: in so-called mixed methods the trial solution satisfies neither the differential equations nor boundary conditions. In Schuleshko's treatment of mixed methods [21], the differential-equation residual is made orthogonal to one set of weighting functions, using (8) as the inner product, while the boundary residual is simultaneously made orthogonal to another set of weighting functions, using an appropriate surface integral as the inner product. If N weighting functions are used, this leads to

*See Mikhlin [11] for a discussion of the contribution by Bubnov in 1913, while his method is the same as the Galerkin method (Mikhlin and others in recent Russian literature call it the Bubnov-Galerkin method), it was Galerkin who developed the method independently of any variational principle.

$2N$ conditions, yet in general only N conditions can be satisfied by the N independent c_i. For this procedure to work, some of the conditions must be discarded, as was noticed by Snyder, Spriggs and Stewart [22] in their discussion of the Galerkin method.

On the other hand, Bolotin [23], Mikhlin [11], and Finlayson [24] have pointed out that for the Galerkin method the dilemma can be resolved by adding the differential-equation residuals to the boundary residuals. The combination is made in such a way that the differential-equation residual, when integrated by parts, cancels identical terms of the boundary residual. The situation is analogous to the treatment of natural boundary conditions in the calculus of variations, and indeed only boundary conditions analogous to natural boundary conditions can be handled in this way. Such a combination of equation and boundary residuals represents a generalization beyond the treatment given by Crandall [1] (page 235), who states that MWR cannot be applied unless the trial solution satisfies all boundary conditions. However, Crandall [1] (page 321) does combine the residuals for eigenvalue problems in which an eigenvalue appears in both the differential equation and boundary conditions. The combination of residuals for more general problems is important in establishing the equivalence between the Galerkin method and several so-called variational methods [25, 24] (see below).

B. Refinements on the Basic Method

Other modifications are possible: Duncan [26] makes the approximating functions satisfy derived (or secondary) boundary conditions which are determined by requiring that the differential equation be satisfied on the boundary. Derived boundary conditions are also used in boundary-layer theory in the von Kármán-Pohlhausen method; other compatibility conditions—such as continuity of the velocity and certain of its derivatives at the edge of the boundary layer—are employed as well. Recently it has been shown [27, 28] that additional compatibility conditions are required to assure good results when the integral method is applied to magnetohydrodynamic boundary-layer problems. In these cases the additional conditions are found by differentiating the differential equation in the direction normal to the surface; all trial solutions must then satisfy this equation at both the solid surface and the edge of the boundary layer. A variation of the collocation method is given by Collatz [3], who differentiates an ordinary differential equation and applies the collocation method to the residual of the resulting equation, too.

Kantorovich and Krylov [29] outline a method for two-dimensional problems in which the residual is required to be zero along a line in the domain (such as $x = y$). The method of reduction to ordinary differential equations, as developed by Kantorovich [30] and independently by Poritsky [31], reduces a partial differential equation to a system of ordinary differential equations. This is the procedure described above for initial-value problems but it can be applied equally well to boundary-value or eigenvalue problems. The spatial averages (Equation (8)) are taken over all the independent variables except one, and the approximate solution is found by solving a set of ordinary differential equations involving this remaining independent variable. While this semi-direct method was originally proposed in the context of variational principles, Kantorovich [32] in 1942 showed its equivalence to the Galerkin method. Even earlier Bickley [12] had applied the Galerkin method to unsteady-state problems in a manner equivalent to the method of reduction to ordinary differential equations. In general, MWR can be used to reduce the number of independent variables in any partial differential equation. The resulting system of equations is simpler (it may be algebraic or ordinary differential equations or even a set of partial differential equations), but its solution remains only an approximate solution to the original problem.

In the collocation method a critical problem is the choice of collocation points. For ordinary differential equations Wright [33] has shown that the residual is minimized if the collocation points are given by the roots of the Chebyshev polynomials.

Naturally the method of weighted residuals can be combined with other methods. Collatz [3] presents a combination of the iteration method and MWR. Yang [34, 35, 36] uses the approximate solution generated by the integral method as the first step in the following procedure for time-dependent problems: the result of the integral method is substituted into those terms involving time-derivatives and equation thereby obtained is solved as a steady-state, nonhomogeneous, partial differential equation. An advantageous coupling of MWR and numerical finite difference methods has been employed by Kaplan [37], Kaplan and Bewick [38], and Kaplan, Marlowe, and Bewick [39] to reduce the computer time necessary to solve certain nuclear reactor problems; the number of independent variables was reduced from four to three or two by using MWR. Other modifications and hybrid schemes are possible and will undoubtedly be proposed as needs arise.

Choice of Approximating Functions

The choice of approximating functions can be crucial in applying MWR. How to arrive at a good, if not the best, selection is an outstanding problem. Certainly any symmetry properties of the system should be exploited but there seems to be no way available at present to do this systematically for all problems. In problems of conventional types it is usually convenient to have the approximating functions satisfy the boundary conditions, and Kantorovich and Krylov [29] show how to construct complete sets of functions which vanish on a boundary of complicated shape. Snyder and Stewart [40] combine this scheme and symmetry arguments to find approximating functions for the velocity vector field in fluid flowing through regularly packed beds of spheres.

Derived boundary conditions can also be used to place restrictions on the approximating functions admitted, and improvement sometimes results [26, 148]. Usually, however, several sets of approximating functions are admissible and it is not possible to choose one as the "best." Heywood and Moffatt [41] even suggest as a qualitative criterion that the approximate

... can be relatively insensitive to different but reasonable choices of the approximating functions.

Methods have been devised for constructing approximating functions especially for eigenvalue problems involving high-order ordinary differential equations of the sort that arise in the theory of convective instability [13, 43, 44]. The approximating functions are just eigenfunctions of one or another lower-order, simpler yet related eigenvalue problem on the same domain. Polynomials are popular approximating functions; they have even been used in cylindrical and spherical domains [45] where proper regard must be taken of possible singularities. Falk [46] uses Hermite polynomials, which are orthogonal on a semi-infinite domain. Other authors [39, 47] emphasize that numerical difficulties (for large N) can be avoided in the Galerkin method if the approximating functions are orthonormalized.

Selecting approximating functions remains somewhat dependent on the user's intuition and expedience, and this is often regarded as a major disadvantage of MWR. Clearly, the question of methods for arriving at optimal choices of approximating functions warrants thorough investigation. Leads may exist in the local solutions and regional expansions used in perturbation methods [47a].

Comparison of Different Criteria

Comparisons of different criteria as applied to the same problem exist only for relatively simple, linear, initial-value problems [12, 25, 48] and boundary-value problems [4, 13, 1]. In the literature on eigenvalue problems the Galerkin method predominates, although there are a few comparisons with the collocation and least-squares methods [1, 4]. The results of these comparisons may be summarized by Crandall's remark [1] (page 375): "The variation between results obtained by applying different criteria to the same trial family... is much less significant than the variations that can result from the choice of different trial families." However, there may be a great difference in the work necessary to obtain the approximate solution when using different criteria. Crandall's experience evidently is based entirely on linear problems. The only comparison for nonlinear problems appears to be the unpublished thesis by Collings [49], as referenced by Ames [2]. Ames comes to the conclusion that the Galerkin method is superior, but cautions that this stand is based on limited experience and may not hold in general [2].

For linear, ordinary differential equations Frazer, Jones and Skan [4] argue that the collocation, least squares, and Galerkin methods are equivalent in the limit as $N \longrightarrow \infty$. Other similarities exist between the methods [1]; e.g. when the approximating functions are chosen to be the eigenfunctions of the linear operator, i.e. $L(u_j) = \lambda u_j$, then the least-squares and Galerkin method coincide.

For self-adjoint (hence necessarily linear) eigenvalue problems, the eigenvalues are real, and Crandall [1] emphasizes that the Galerkin method leads to symmetric matrices in Equation (15)—and hence real-valued approximations—whereas the other methods may give complex eigenvalues as approximations to the exact real eigenvalues. The least-squares method is particularly unsuited for linear eigenvalue problems because it turns the linear problem into a nonlinear one [1].

The least-squares procedure for eigenvalue problems as outlined by Becker [50] differs somewhat from that of Crandall [1]; Becker does not have such a difficulty in the first approximation. Whereas Crandall uses the weighting function $\partial I/\partial c_1$ for the first approximation, Becker uses $\partial I/\partial \lambda$, where λ is the eigenvalue. Consider the linear eigenvalue problem

$$L(u) + \lambda u = 0 \tag{20}$$

with $u = 0$ on the boundary. For the first approximation with a trial solution $u^* = c_1 u_1$, the residual is

$$R(u_1) = c_1(L(u_1) + \lambda u_1) \tag{21}$$

The mean square error is then

$$I = c_1^2 \int_V [L(u_1) + \lambda u_1]^2 \, dV \tag{22}$$

Crandall apparently would determine λ from

$$\partial I/\partial c_1 = 0 = 2c_1 \int_V [L(u_1) + \lambda u_1]^2 \, dV \tag{23}$$

which is a quadratic in λ and may lead to complex values of λ. Becker, on the other hand, would determine λ from

$$\partial I/\partial \lambda = 0 = 2c_1^2 \int_V [L(u_1) + \lambda u_1] u_1 dV \tag{24}$$

which is linear in λ and gives real values as long as the equation and u_1 are real. Becker's procedure appears to be simpler for the first approximation.

For higher approximations both procedures lead to nonlinear equations for this linear problem. Crandall would use as weighting functions $\partial I/\partial c_i$, $i = 1, 2, \ldots, N$, and Becker would use $\partial I/\partial \lambda$, $\partial I/\partial c_j$, $j = 2, 3, \ldots, N$. The latter is thus using the eigenvalue λ as one of the parameters and is also exploiting the fact that the mean square error can be minimized as a function of $\{\alpha_j \equiv c_j/c_1\}$, rather than $\{c_j\}$ since

$$I(c_1, c_2, \ldots, c_N, \lambda) = c_1^2 I(1, \alpha_2, \ldots, \alpha_N, \lambda). \tag{25}$$

For initial-value problems, the least-squares method must be applied carefully and has certain disadvantages. The method is applicable if the time dependence of the approximate solution is specified—in other words, semi-direct methods cannot be used in the method of least squares. Consider the problem $\partial u/\partial t = L(u)$ and assume a trial solution of the form $u^* = u_s + \sum a_i u_i(x, t)$. Then the functional I, representing the mean square residual, can be minimized:

$$I = \int_0^T \int_V \left[\frac{\partial u}{\partial t} - L(u)\right]^2 dV \, dt. \tag{26}$$

Of course the solution depends on the value of T. If the upper limit of integration is infinite, the solution may no longer have this ambiguity. This was the ap-

approach taken by Buckley [12] in his least-squares calculations for time-dependent problems.

Oftentimes, however, the time dependence of the solution is difficult to guess and the trial solution must involve undetermined functions of time, $u^* = u_s + \sum_{i=1}^{N} c_i(t)\, u_i(\mathbf{x})$. The mean square residual is

$$I = \int_v \left[\frac{\partial u}{\partial t} - L(u)\right]^2 dV \tag{27}$$

Now, however, I depends on time and involves time derivatives. Consequently it cannot in general be made a minimum for all time by any set of functions $c_i(t)$; this was shown by Citron [51] and Finlayson and Scriven [25]. Consequently, if a semi-direct method is used to solve this type of problem the term least squares is a misnomer because the mean square residual is not being minimized.

The least-squares method is discussed at length in a monograph by Becker [50]. Listing criteria which he maintains a good variational method must satisfy, he concludes that the least-squares method is the best general criterion of MWR. Becker's list includes the following points: (i) errors should be minimized in some sense; (ii) the functional should be positive definite; (iii) the procedure should be capable of treating initial-value problems, as well as others. These seem somewhat slanted toward the least-squares method; indeed, items (i) and (ii) cannot be realized for all problems except in the least-squares method. Yet no one has shown that a solution is necessarily best because its mean square residual is smallest; such a definition or proof will certainly depend on the particular application. In addition, the least-squares method can be used to treat initial-value problems in only a limited way, as shown above. Furthermore, an important point should be added to the list of desiderata—the method should be simple to apply. As already shown, this criterion immediately eliminates the least-squares method for linear eigenvalue problems because it turns a linear problem into a more difficult nonlinear one. Becker realized that his conclusion may not always be valid [50] (page 61): "While the least-squares method seems to be the most suitable general approach, in specific applications (in which some specific criteria may be added to our 'general' list) other methods may be preferable." Becker illustrates the advantages of the method of least-squares by solving a set of nonlinear, time-dependent partial differential equations which model the fuel depletion in a nuclear reactor; he finds results that compare well with the more lengthy numerical solutions.

In this discussion of the various criteria of MWR, the Galerkin method has been distinguished from the method of moments by means of the weighting functions used in the two. In the Galerkin method, the weighting functions must be the same set of functions which are used for the trial solution, whereas in the method of moments the weighting functions can be some other set of functions. This distinction is not always made [52] and is probably unimportant in practice, although the two methods have different histories and may have different convergence properties. There are inconsistencies of terminology in the literature; for example, Kawaguit [53] used the method of moments rather than the Galerkin method as he claimed, for the weighting functions differ from the approximating functions in his work. Another example of confusing nomenclature is the name method of integral relations, which refers to a generalization of the subdomain method; it is adequately reviewed by Belotserkovskii and Chushkin [54].

C. Convergence Theorems

Galerkin Method

After introduction of the Galerkin method in 1915 some twenty-five years elapsed before the convergence of the method was studied. Even today much remains to be done; only a few theorems have been proved, and these pertain exclusively to linear problems. Repman [55] was the first to prove convergence of solutions obtained by the Galerkin method though only for a certain Fredholm-type integral equation. Petrov [56] then studied the convergence of the Galerkin method for eigenvalue problems of fourth-order ordinary differential equations—in particular, the Orr-Sommerfeld equation of hydrodynamic stability theory. Keldysh [57] treated general ordinary differential equations and also second-order elliptic partial differential equations. Mikhlin [11] later simplified Keldysh's proofs. The equations are of the form

$$L(u) \equiv -\sum_{i,\,j=1}^{m} \frac{\partial}{\partial x_i}\left(A_{ij}\frac{\partial u}{\partial x_j}\right) + \sum_{i=1}^{m} B_i \frac{\partial u}{\partial x_i} + Cu = f \tag{28}$$

Both Keldysh and Mikhlin prove that the first derivatives of the Galerkin approximate solution converge in the mean to the first derivatives of the exact solution.

Whenever the Rayleigh-Ritz and Galerkin methods coincide (see below), the convergence proofs for the Rayleigh-Ritz method imply convergence of the Galerkin method, too. Thus the Galerkin-convergence proofs given by Kantorovich and Krylov [29] apply only to specific problems with a minimum or maximum principle, whereas the convergence proofs mentioned here are applicable to problems whether or not they have a corresponding variational principle. It has been claimed [40] that completeness of the set of approximating functions is sufficient to assure convergence, but the proofs given by Mikhlin and others show clearly that this is not enough.

Recently convergence proofs have become available for certain eigenvalue problems associated with hydrodynamic stability investigations [58, 15].

Results applicable to unsteady-state problems are less extensive. Faedo [59] applied the Galerkin method to a hyperbolic differential equation and inspired the important work of Green [60], who proved the uniform convergence of the Galerkin method when ap-

plied to the following equation

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial t} - g(x, t)u = f(x, t) \qquad (29)$$

The Galerkin method has been used to prove the existence of weak solutions to (i) the Navier-Stokes equations with time dependence [61], and (ii) equations representing the unsteady-state transport equation with a known velocity field [62]:

$$\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(A_{ij}\frac{\partial u}{\partial x_j}\right) + \sum_{i=1}^{m} B_i \frac{\partial u}{\partial x_i} + Cu - \frac{\partial u}{\partial t} = F(x, t) \qquad (30)$$

Recently the Galerkin method has been applied to the Taylor problem with time-dependent disturbances [63], and a method has been developed to generate improvable, pointwise upper and lower bounds—and hence error bounds—for the solution to Equation (30) [129a].

Other Methods

Convergence proofs are rarely available for the other criteria of MWR. The notable exception is the least-squares method, which is well-treated (for boundary-problems) in the notable text by Mikhlin [11]. Mikhlin proves conditions which insure that the method of least squares gives a sequence of approximate solutions which converge in the mean to the exact solution. Furthermore, the mean-square-error of the approximate solution can be determined. He also points out that the least-squares method converges more slowly than the Ritz method (when the latter can be applied) but may give uniform convergence rather than convergence in the mean. Some results for the collocation method are given by Kadner [64], while the method of moments is treated by Kravchuk [65].

Nonlinear Problems

Very little is known about the convergence of MWR for nonlinear problems without a corresponding variational principle. Krasnosel'ski [66] presents theorems—mostly without proofs—for the Galerkin method applied to nonlinear integral equations. Glansdorff [67] mentions a forthcoming proof of the convergence of the local potential method, which is identical in application to the Galerkin method; he treats the steady-state heat conduction equation with temperature-dependent thermal conductivity. Of course, as Ames [2] has pointed out, convergence proofs are not as useful as error bounds. Even a computer does not make it possible to calculate infinitely many terms and when truncating the series one always wonders how good the resulting approximate solution is. Comparison of successive approximations is an aid in such a case, but even an approximate solution that seems to be converging may not be converging to the exact solution. The available convergence theorems and error bounds are so scarce that engineer and applied scientist must usually extrapolate from previously tested results for other problems to new situations when applying approximate methods.

D. Comparison to Other Methods

Separation of Variables

The Galerkin method is related to a wide variety of other approximate methods as well as to some exact methods of analysis. In particular, it can be shown [29, 22, 24] that if a problem yields to the method of separation of variables and if the Galerkin method is applied in a certain way*, then the two solutions are the same, provided the Galerkin method is carried through to completion. Of course in numerical calculations, after obtaining an exact solution in the form of an infinite series, one calculates only a finite number of terms as a matter of practical necessity.

Variational Methods

There is also a close relationship between the Galerkin method and the Ritz or Rayleigh-Ritz† method when the latter can be applied [14, 69, 70, 1, 21, 3, 29, 71, 72]. In particular, if the same trial functions are used, the resulting calculations are identical. Contrary to a currently prevalent opinion, this equivalence still persists when the trial functions do not satisfy the natural boundary conditions [23, 11, 25, 24], which they need not do in the Rayleigh-Ritz method. The boundary residual is either added or subtracted to the differential equation residual, and the calculations are again equivalent to the Ritz or Rayleigh-Ritz method. The choice of adding or subtracting is dictated either by mathematical convenience—part of the differential-equation residual can be integrated by parts to cancel part of the boundary residual—or by the physics—the differential equation and boundary conditions both come from macroscopic balances taken over the volume and surface, respectively; these macroscopic balances can be combined in only one way, and the residuals are combined in exactly the same way. A very important difference between the Rayleigh-Ritz method and the Galerkin method is that in the former some functional—possibly representing an eigenvalue—is being minimized or maximized. Consequently the approximate values of the functional represent either upper or lower bounds. In the Galerkin method this information is missing; exactly the same values would be obtained, but one would not know that these were upper or lower bounds. However, when the variational integral is of no significance, the Galerkin method, because of its generality, may be preferred. The variational and Galerkin methods are compared schematically in Figure 1.

Most variational principles are merely stationary principles, rather than minimum or maximum principles.

*The approximating functions in the Galerkin method must be the eigenfunctions found by the separation of variables and the Galerkin method must be applied to the initial conditions as well as to the differential equation. Such a result means simply that if the exact solution is contained in the trial function, the Galerkin method will find it.

†Though there is basically but a single method, it is convenient to follow the custom (scarcely universal) of distinguishing between the "Rayleigh-Ritz method" when it is applied to minimum or maximum principles and the "Ritz method" when it is applied to merely stationary principles.

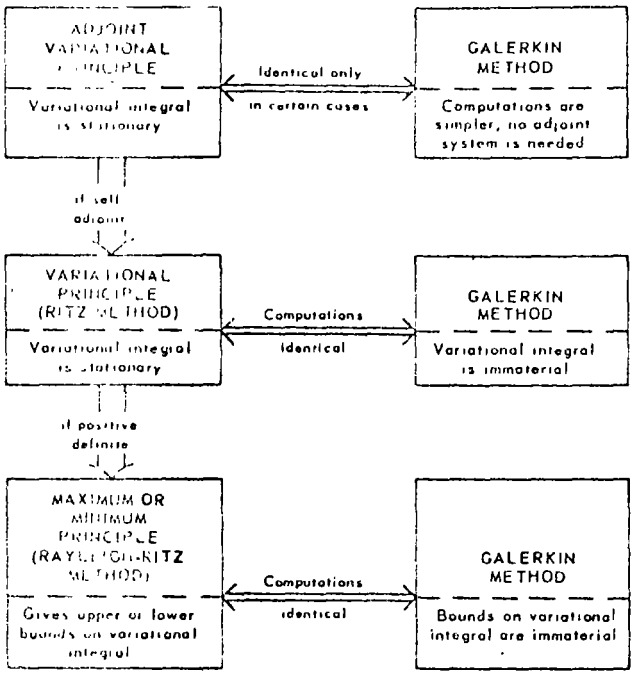


Figure 1. Comparison of Variational and Galerkin Methods for Linear Problems.

In such cases, the Ritz method is again equivalent to the Galerkin method. The calculations are identical; the results are identical; but in the variational method one knows that the variational integral is being made stationary, i.e., insensitive to changes in the trial solution. If the variational integral has physical significance and is the quantity of interest, then the variational methods have an advantage over the Galerkin method even though the answers are the same.

Adjoint Variational Methods

Variational principles exist for linear problems only if they are self-adjoint. For non-self-adjoint linear problems, variational principles can be formulated for the original equations and their adjoints, and again MWR is related to the corresponding variational methods. The impetus for using the adjoint operator in variational formulations seems to stem from Morse and Feshbach [73], who gave a variational principle for the unsteady-state heat conduction equation; Roussopoulos [74] also gave a variational principle for any linear non-self-adjoint problem. Schmit [75] and Washizu [76] have applied such a principle to the unsteady-state heat conduction equation, while Selengut [77, 78] developed the idea for nuclear reactor problems. Many other examples exist, such as those of Nichols and Bankoff [79] for convective diffusion of heat, Finlayson [24] for convective diffusion of a multicomponent mixture of chemical species; Lewins [80], Slattery [81], Flumerfelt and Slattery [82] for extensions to nonlinear problems; and many authors [83–94, 38] for nuclear reactor and associated problems. In applications to linear non-self-adjoint problems, the method of weighted residuals yields the same results as any of these variational principles as long as the weighting functions for the original equations are taken as the approximating functions for the adjoint, and vice versa [75, 24]. The question then arises as to whether this variational method, which can be regarded as an application of one form of MWR, is preferable to Galerkin's method, which does not require the complication of an auxiliary adjoint system. There is some evidence that the adjoint variational principle leads to slightly better results [95–97], and Clark and Hansen [98] imply that the use of adjoint weighting functions might speed convergence. Kaplan and Bewick [38] claim that the variational method is the best strategy in that it gives better answers more often. However, they go on to say:

> "Of course, practical considerations may introduce yet another meaning of the work "best"; namely, "most economical." In this sense we find that the Galerkin method (which uses the trial functions also as weighting functions) is, in most instances, preferable to the variational method, since it gives results which are almost as good but does not require separate calculations of the weighting functions."

In essence, the adjoint variational method trades increased complexity for possibly better results: there still is no clear-cut answer to the question of whether the Galerkin method or the adjoint-variational method is best.

For certain initial-value problems there may be no difference between the variational method and Galerkin's method if the semi-direct approach is used in both. Whenever the corresponding steady-state problem is self-adjoint, it is reasonable to expand the unsteady-state solution and its adjoint in terms of the same functions of position with unknown functions of time as coefficients:

$$u = u_s + \sum_{i=1}^{N} c_i(t) u_i(\mathbf{x}) \tag{31}$$

$$u^* = u_s + \sum_{i=1}^{N} c_i^*(t) u_i(\mathbf{x}) \tag{32}$$

Because the approximating functions are the same for both u and u^*, the weighting functions in the adjoint method are the same as those in the Galerkin method. Consequently, the solutions are identical, whether derived by the adjoint method or the more direct Galerkin method.

The adjoint system is also useful for eigenvalue problems. Roberts [99] presents the general theory, and examples can be found in the works of Chandrasekhar [43] and DiPrima [100] as well as others. While introduction of the adjoint system does increase the complexity of the problem—particularly the boundary conditions—some advantage is gained over the straightforward application of MWR because the eigenvalue is made stationary, and hence insensitive to changes in the trial function. This advantage does not usually apply to boundary and initial-value problems since the variational integral is seldom of interest in those cases.

Method of Least Squares

Mikhlin [11] points out that for boundary-value problems the least-squares method for

$$Lu = f \quad \text{in } V \tag{33}$$

$$B_i u = 0 \quad \text{on } S \tag{34}$$

is equivalent to applying the calculus of variations to the equation

$$L^*(Lu - f) = 0 \tag{35}$$

Consider the following minimum principle: minimize the functional

$$I = \int_V (Lu - f)^2 \, dV + \int_S (B_i u)^2 \, dS \tag{36}$$

among all functions u having the appropriate continuity and differentiability requirements. The natural boundary conditions corresponding to this variational principle are of the form [50]

$$N_i(Lu - f) = 0 \tag{37}$$

where N_i are differential operators. The Euler equation is just equation (35) and the equivalence with equation (33) rests with the premise that the equation

$$L^* u = 0 \tag{38}$$

$$B_i u = 0 \tag{39}$$

has only the trivial solution. Note that the natural boundary conditions (37) are similar to the compatibility conditions mentioned above in connection with the integral method for magnetohydrodynamic boundary-layer problems.

Method of the Local Potential

A procedure based on the so-called local potential of Prigogine and Glansdorff [101–104] has been proposed as a variational method for determining approximate solutions to boundary-value [105–108], eigenvalue [109], and more recently initial-value problems [108]. Rosen [110–113] used the same type of computational scheme earlier. The actual applications of these methods have been shown to be equivalent to the Galerkin method [114, 24]. Moreover, it has been demonstrated that the variational integral is not stationary in the local potential method and that no minimum principle exists in applications [114, 24]. Consequently, the advantages usually associated with variational principles are missing from the local potential method, which can be regarded as a disguised application of the Galerkin method. See Ref. 129b for a more detailed critique.

Lagrangian Thermodynamics

The so-called variational methods due to Biot [115–121] and others [122–128, 51] are also equivalent to the Galerkin method [25, 24]. In these Lagrangian thermodynamic methods there is no variational integral which is being made stationary [25, 79]; their sole significance appears to be as means for generating a computational scheme. That scheme is, however, identical to the Galerkin method, which is more straightforward and applicable to a broader range of situations. There is no reason that the Galerkin method should not be preferred, so far as the authors know. See Ref. 129b for a more detailed critique.

E. Applications

The general features of MWR in its numerous versions and various refinements have been presented, and its relationships to certain other approximation methods have been sketched. Which of all these methods are superior, and over just what ranges of circumstances the superiority exists, are matters that can be settled finally only on the basis of representative applications. More systematic comparative studies and evaluations are needed than have been reported to date. Until they are forthcoming the investigator of a new problem can expect little more help than he can get out of seeing how others have handled more or less similar problems. References 130–187 have been selected as much to illustrate pitfalls, shortcomings, and failures as to cite the attractive features and successes of different versions of MWR. The preponderance of recent papers accurately reflects the upsurge of applications of these methods in one field after another; the emphasis on problems of flow and transport is conditioned by interests of the authors. The popularity of the integral methods which originated in boundary-layer studies [18, 20, 34–36, 136–153, 160–163, 166–169] can be discounted in part as a tradition perpetuated by formal instruction beginning with elementary texts in fluid mechanics and heat transfer.

Beyond any guidance he can get from past experience the problem-solver can look for reassurances in comparisons of different forms of trial solutions and of successive approximations in any one form: the appearances of convergence with more numerous adjustable parameters and of insensitivity to form of approximating functions do lend confidence to results. So do close matches with established information on special cases and limiting cases. It is also true that MWR, like variational methods, may yield better estimates of properties of the solution at large, such as an integral or eigenvalue, than of the solution itself. The main advantage and disadvantage of MWR are contained in the same feature, namely, that the results depend on more or less arbitrary decisions by the user. Intuition, experience, any available information all can be rapidly exploited but the reliability of the results is frequently hazy. Hopefully this review sheds light on the basic issues and will be useful to those interested in applying weighted residual methods and related techniques.

F. Areas for Further Research

Of the unsolved problems concerning MWR the following are most important in the opinion of the authors:

(1) Choice of criterion in MWR. Systematic comparative studies using representative (nonlinear) problems are needed. The least-squares procedure for nonlinear problems particularly warrants attention. The Galerkin method and adjoint variational method for linear problems need to be compared.

(2) Development of rational methods for selecting approximating functions, for example by systematically applying symmetry properties of the problem. A related question is whether linear combinations of functions are more effective than nonlinear forms, especially for nonlinear problems.

(3) Definition of mathematical or engineering criteria for identifying optimal approaches in (1) and (2). Optimization theory should be brought to bear on the decisions that have to be made by users of weighted-residual and other approximation methods.

(4) Derivation of error bounds for approximate solutions by MWR. And though they are not as directly useful, convergence proofs are needed, especially for nonlinear problems.

Acknowledgment

This review stems from research supported in part by the National Science Foundation through a Graduate Fellowship received by B. A. Finlayson and in part by the Air Force Office of Scientific Research through AFOSR Grant Number 219-63.

References

1 Crandall, S. H., "Engineering analysis," McGraw-Hill, 1956, AMR 12(1959), Rev. 1122.

2 Ames, W. F., "Nonlinear partial differential equations in engineering," Academic Press, 1965; AMR 19(1966), Rev. 4724.

3 Collatz, L., "The numerical treatment of differential equations," Springer, 1960; AMR 14(1961), Rev. 1738.

4 Frazer, R. A., Jones, W. P. and Skan, S. W., "Approximations to functions and to the solutions of differential equations," Gt. Brit. Aero. Res. Council Rept. and Memo. 1799 (1937); reprinted in Gt. Brit. Air Ministry Aero. Res. Comm. Tech. Rept. 1, 517-549 (1937).

5 Biezeno, C. B. and Koch, J. J., Over een nieuene methode ter berekening van vlokke platen met toepassing op enkele voor de techniek belangrijke belastingsgevallen, *Ing. Grav.* 38, 25-36(1923).

6 Biezeno, C. B., Over een vereenvoudiging en over een uitbreiding van de methode van Ritz, *Christiaan Huygens* 3, 69 (1923-1924).

7 Biezeno, C. B., Graphical and numerical methods for solving stress problems, in Proc. First Intern. Congr. Appl. Mech., Delft 1924, pp. 3-17.

8 Courant, R., Remark on "weighted averages of the residual" in discussion following Biezeno's paper in Proc. First Intern. Congr. Appl. Mech. Delft 1924, pg. 17.

9 Picone, M., Sul metodo delle minime potenze ponderate e sul metodo di Ritz per il calcolo approssimato nei problemi della fisica-matematica, *Rend. Circ. Mat. Palermo* 52, 225-253 (1928).

10 Galerkin, B. G., Sterzhni i plastiny. Ryady v nekotorykh voprosakh uprogogo ravnovesiya sterzhnei i plastin (Rods and plates. Series occurring in some problems of elastic equilibrium of rods and plates), *Vestn. Inzhen. i Tekh. Petrograd* 19, 897-908 (1915); Translation 63-18924, Clearinghouse Fed. Sci. Tech. Info.

See also: On the seventieth anniversary of the birth of B. G. Galerkin (in English with a list of his publications), *Prikl. Mat. Mekh.* 5, 337-341 (1941).

11 Mikhlin, S. G., "Variational methods in mathematical physics," Pergamon/Macmillan, 1964.

12 Bickley, W. G., Experiments in approximating to the solution of a partial differential equation, *Phil. Mag.* (7) 32, 50-56(1941).

13 Ertugen, A. C., Transverse impact on beams and plates, *J. Appl. Mech.* 20, 461-468 (1953).

14 Duncan, W. J., "The principles of Galerkin's method," Gt. Brit. Aero. Res. Council Rept. and Memo. 1848 (1938); reprinted in Gt. Brit. Air Ministry Aero. Res. Comm. Tech. Rept. 2, 589-612(1938).

15 Sani, R. L., "Convergence of a generalized Galerkin method for certain fluid stability problems," Rensselaer Polytechnic Inst. Math. Rept. no. 63 (1964).

16 Yamada, H., *Rept. Res. Inst. Fluid Engng., Kyushu Univ.* 3, 29 (1947); referred to by H. Fujita in Ref. 129. Also: A method of approximate integration of the laminar boundary-layer equation, *Rept. Res. Inst. Fluid Engng., Kyushu Univ.* 6, 87-98 (1950); AMR 5(1952), Rev. 488.

17 Kravchuk, M. F., Application of the method of moments to the solution of linear differential and integral equations (in Ukrainian), *Kiev. Soobshch. Akad. Nauk UkSSR* 1, 168(1932); referred to by P. Shuleshko in Ref. 21 and L. V. Kantorovich and V. I. Krylov in Ref. 29. Another citation is given by S. G. Mikhlin in Ref. 11.

18 Pohlhausen, K., Zur näherungsweisen Integration der Differential-gleichung der laminar en Grenzschicht, *ZAMM* 1, 252-268(1921).

19 Von Kármán, Th., Über laminare und turbulente Reibung, *ZAMM* 1, 233-252 (1921); NACA Tech. Memo. 1092 (1946).

20 Goodman, T. R., Application of integral methods to transient nonlinear heat transfer, in "Adv. Heat Transfer" 1, 51-122(1964); AMR 18(1965), Rev. 3009.

21 Shuleshko, P., A new method of solving boundary-value problems of mathematical physics, Generalizations of previous methods of solving boundary-value problems of mathematical physics, *Austral. J. Appl. Sci.* 10, 1-7, 8-16 (1959); AMR 13(1960), Rev. 3.

22 Snyder, L. J., Spriggs, T. W. and Stewart, W. E., Solution of the equations of change by Galerkin's method, *AIChE J.* 10, 535-540 (1964).

23 Bolotin, V. V., "Nonconservative Problems of the Theory of Elastic Stability," Pergamon/Macmillan, 1963, pg. 60; AMR 16(1963), Rev. 46.

24 Finlayson, B. A., "Approximate solutions of equations of change. Convective instability by active stress," Ph.D. thesis, Univ. of Minnesota, 1965.

25 Finlayson, B. A. and Scriven, L. E., The method of weighted residuals and its relation to certain variational principles for the analysis of transport processes, *Chem. Engng. Sci.* 20, 395-404 (1965).

26 Duncan, W. J., "Galerkin's method in mechanics and differential equations," Gt. Brit. Aero. Res. Council Rept. and Memo. 1798 (1937); reprinted in Gt. Brit. Air Ministry Aero. Res. Comm. Tech. Rept. 1, 484-516 (1937).

27 Hugelman, R. D. and Haworth, D. R., An MHD boundary-layer compatibility condition, *AIAA J.* 3, 1367-1369 (1965); AMR 19(1966), Rev. 2539.

28 Hugelman, R. D., Application of Pohlhausen's method to stagnation-point flow, *AIAA J.* 3, 2158-2159 (1965); AMR 19(1966), Rev. 3014.

29 Kantorovich, L. V. and Krylov, V. I., "Approximate methods in higher analysis," Interscience, 1958; AMR 12(1959), Rev. 4830.

30 Kantorovich, L. V., A direct method of solving the problem of the minimum of a double integral (in Russian), *Izv. Akad. Nauk SSSR* 5, 647-652 (1933).

31 Poritzky, H., "The reduction of the solution of certain partial differential equations to ordinary differential equations," in Trans. Fifth Intern. Congr. Appl. Mech., Cambridge, Mass., 1938, pp. 700-707.

32 Kantorovich, L. V., Application of Galerkin's method to the so-called procedure of reduction to ordinary differential equations (in Russian with English summary), *Prikl. Mat. Mekh.* 6, 31-40 (1942).

33 Wright, K., Chebyshev collocation methods for ordinary differential equations, *Comput. J.* 6, 358-365 (1964).

34 Yang, K. T., "Calculation of unsteady heat conduction in a single layer and composite finite slabs with and without property variations by an improved integral procedure," in Intern. Developments in Heat Transfer, A.S.M.E. (1961), pp. 18-27.

35 Yang, K. T., An improved integral procedure for compressible laminar boundary layer analysis, *J. Appl. Mech.* 28, 9-20 (1961); AMR 14(1961), Rev. 6188.

Yang, K. T., Laminar forced convection of liquids in tubes with variable viscosity, *Trans. ASME, J. Heat Transfer* 84, 353–362 (1962).

Kaplan, S., Some new methods of flux synthesis, *Nucl. Sci. Engng.* 13, 22–31 (1962).

Kaplan, S. and Bewick, J. A. "Space and time synthesis by the variational method." Bettis Techn. Review, Reactor Technology, 27–44, WAPD-BT-28, (1963), Clearinghouse Fed. Sci. Tech. Info.

Kaplan, S., Marlowe, O. J., and Bewick, J., Application of synthesis techniques to problems involving time dependence, *Nucl. Sci. Engng.* 18, 163–176 (1964).

Snyder, L. J. and Stewart, W. E., Velocity and pressure profiles for Newtonian creeping flow in a regular packed bed of spheres, *AIChE J.* 12, 167–173 (1966); AMR 19(1966), Rev. 6002.

Heywood, J. B. and Moffatt, W. C., Validity of integral methods in MHD boundary layer analyses, *AIAA J.* 3, 1565–1567 (1965).

Reid, W. H. and Harris, D. H., On orthogonal functions which satisfy four boundary conditions, *Astrophys. J., Suppl.* 3, 429–452 (1958).

Chandrasekhar, S., "Hydrodynamic and hydromagnetic stability," Clarendon Press, 1961.

Krueger, E. R. and DiPrima, R. C., The stability of a viscous fluid between rotating cylinders with an axial flow, *J. Fluid Mech.* 19, 528–538 (1964).

Lardner, T. J. and Pohle, F. V., Application of the heat-balance integral to problems of cylindrical geometry, *J. Appl. Mech.* 28, 310–312 (1961); AMR 15(1962), Rev. 2860.

Falk, S., Das Verfahren von Rayleigh-Ritz mit hermiteschen Interpolationspolynomen, *ZAMM* 43, 149–166 (1963).

Caldwell, W. R., Kaplan, S. and Marlowe, O. J., "Equations and programs for synthesis approximations," WAPD-TM-377 (1963), Clearinghouse Fed. Sci. Tech. Info.

Van Dyke, M., "Perturbation methods in fluid mechanics," Academic Press, 1964; AMR 19(1966), Rev. 1710.

Thorsen, R. and Landis, F., Integral methods in transient heat conduction problems with nonuniform initial conditions, *Internat. J. Heat Mass Transfer* 8, 189–192 (1965); AMR 19(1966), Rev. 3145.

Collings, W. Z., "The method of undetermined functions as applied to nonlinear diffusion problems," M.M.E. thesis, Univ. of Delaware, 1962.

Becker, M., "The principles and applications of variational methods," Res. Monogr. no. 27, MIT Press, 1964.

Citron, S. J. A note on the relation of Biot's method in heat conduction to a least-squares procedure, *J. Aero/space Sci.* 27, 317–318 (1960); AMR 13(1960), Rev. 5951.

Hildebrand, F. B., "Methods of applied mathematics," Prentice-Hall, 1954.

Kawaguti, M., The critical Reynolds number for the flow past a sphere, *Rept. Inst. Sci. Technol. Tokyo* 2, 66 (1948), English version in *J. Phys. Soc. Japan* 10, 694–699 (1955).

Belotserkovskii, O. M. and Chushkin, P. I., "The numerical method of integral relations," NASA Tech. Transl. F-8356 (1964); AMR 18(1965), Rev. 2578.

Repman, Yu. V., A problem in the mathematical foundations of Galerkin's method for solving problems on the stability of elastic systems (in Russian with French summary), *Prikl. Mat. Mekh.* 4, 2, 3–6 (1940).

Petrov, G. I., Application of Galerkin's method to a problem involving steady flow of a viscous fluid, (in Russian with English summary), *Prikl. Mat. Mekh.* 4, 3, 3–12 (1940).

Keldysh, M. V., On B. G. Galerkin's method for the solution of boundary-value problems," *Izv. Akad. Nauk SSSR, Ser. Mat.* 6, 309–330 (1942); NASA Tech. Transl. F-195 (1964).

DiPrima, R. C. and Sani, R. L., The convergence of the Galerkin method for the Taylor-Dean stability problem, *Quart. Appl. Math.* 23, 183–187 (1965).

Faedo, S., Un nuovo metodo per l'analisi esistenziale e quantitativa dei problemi di propagazioni, *Ann. Scuola Normale Superiore di Pisa* (3) 1, 1–41 (1949).

Green, J. W., An expansion method for parabolic partial differential equations, *J. Res., Natl. Bur. Stds. (U.S.)* 51, 127–132 (1953).

Hopf, E., Über die Anfangswertaufgabe für die hydrodynamschen Grundgleichungen, *Math. Nachr.* 4, 213–231 (1950/1951).

Il'in, A. M., Kalashnikov, A. S. and Oleinik, O. A., Second-order linear equations of parabolic type, *Russian Math. Surveys* 17, 1–144 (1962).

Meister, B., Die Anfangswertaufgabe für die Störungsdifferentialgleichungen des Taylorschen Stabilitätsproblem, *Arch. Rational Mech. Anal.* 14, 81–107 (1963); AMR 17(1964), Rev. 3964.

Kadner, E., Untersuchungen zur Kollokationsmethode, *ZAMM* 40, 99–113 (1960); AMR 14(1961), Rev. 1172.

Kravchuk, M. F., On the convergence of the method of moments for partial differential equations (in Ukrainian), *Zb. Ins-ta Mat., Akad. Nauk UkSSR* 1, 23–27 (1936); referred to by L. V. Kantorovich and V. I. Krylov in Ref. 29.

Krasnosel'ski, M. A., "Topological methods in the theory of nonlinear integral equations," Pergamon/Macmillan 1964.

Glansdorff, P., "Mechanical flow processes and variational methods based on fluctuation theory," pp. 45–54 in Ref. 68.

Donnelly, R. J., Herman, R. and Prigogine, I., editors, "Non-equilibrium thermodynamics, variational techniques and stability," Univ. of Chicago Press, 1966.

Biezeno, C. B. and Grammel, R., "Engineering dynamics, Vol. I: Theory of elasticity," Blackie and Son, 1955, pg. 173.

Sokolnikoff, I. S., "Mathematical theory of elasticity," McGraw-Hill, 1946, pg. 413.

Singer, J., On the equivalence of the Galerkin and Rayleigh-Ritz methods, *J. Roy. Aero. Soc.* 66, 592 (1962); AMR 16(1963), Rev. 1321.

Reiss, E. L., Supplement bound with "Problems of mathematical physics" by Lebedev, N. N., et al, Prentice-Hall, 1965, entitled "Kantorovich's Method," pp. 401–413.

Morse, P. M. and Feshbach, H., "Methods of theoretical physics," McGraw-Hill, 1953, Chapt. 3.

Roussopoulos, P., Méthodes variationnelles en théorie des collisions, *Compte rendus* 236, 1858–1860 (1953).

Schmit, L. A., "Application of the variational method, the Galerkin technique, and normal coordinates in a transient temperature distribution problem," WADC Tech. Rept. 56-287 (1956); AD-97326, Clearinghouse Fed. Sci. Tech. Info.

Washizu, K., "Application of the variational method to transient heat conduction problems," MIT Tech. Rept. 25–17 (1955); AD-59951, Clearinghouse Fed. Sci. Tech. Info.

Selengut, D. S., "Variational analysis of multidimensional systems," in Nucl. Physics Res. Quart. Rept. for Oct.-Dec., 89–124 (1958); HW-59126 (20 Jan. 1959).

Selengut, D. S., On the derivation of a variational principle for linear systems, *Nucl. Sci. Engng.* 17, 310–311 (1963).

Nichols, R. A. and Bankoff, S. G., Adjoint variational principles for convective diffusion, *Internat. J. Heat Mass Transfer* 8, 329–335 (1965); AMR 18(1965), Rev. 5880.

Lewins, J., A variational principle for nonlinear systems, *Nucl. Sci. Engng.* 12, 10–14 (1962).

Slattery, J. C., A widely applicable type of variational integral. I, *Chem. Engng. Sci.* 19, 801–806 (1964).

Flumerfelt, R. W. and Slattery, J. C., A widely applicable type of variational integral. II, *Chem. Engng. Sci.* 20, 157–163 (1965).

Lippmann, B. A. and Schwinger, J., Variational principles for scattering processes. I," *Phys. Rev.* 79, 469–480 (1950).

Corngold, N., Slowing down of neutrons in infinite homogeneous media, *Proc. Phys. Soc. London* A70, 793–801 (1957).

Lewins, J., The time-dependent importance of neutrons and precursors, *Nucl. Sci. Engng.* 7, 268–274 (1960).

Lewins, J., Variational representations in reactor physics derived from a physical principle, *Nucl. Sci. Engng.* 8, 95–104 (1960).

Dougherty, D. E. and Shen, C. N., The space-time

neutron kinetic equations obtained by the semi-direct variational method, *Nucl. Sci. Engng.* 13, 111–118 (1962).

88 Goldstein, R. and Cohen, E. R., Theory of resonance absorption of neutrons, *Nucl. Sci. Engng.* 13, 132–140 (1962).

89 Pomraning, G. C. and Clark, M., Jr., The variational method applied to the monoenergetic Boltzmann equation. I and II, *Nucl. Sci. Engng.* 16, 147–154, 155–164 (1963).

90 Lewins, J., The weighting functions of reactor physics, *J. Brit. Nucl. Energy Soc.* 2, 326–334 (1963).

91 Gast, P. F., A variational principle for heterogeneous resonance capture as a basis for approximation methods, *Nucl. Sci. Engng.* 19, 196–202 (1964).

92 Mika, J. R., On the variational method applied to the monoenergetic Boltzman equation, *Nucl. Sci. Engng.* 19, 377–378 (1964).

93 Lewins, J., Time-dependent variational principles for nonconservative systems, *Nucl. Sci. Engng.* 20, 517–520 (1964).

94 Lewins, J., "Importance: the adjoint function," Pergamon, 1965.

95 Fischer, P. G., Group parameters for few-group perturbation calculations, *Trans. Amer. Nucl. Soc.* 2, no. 2, 164 (1959); Summary 16-4.

96 Calame, G. P. and Federighi, F. D., A variational procedure for determining spatially dependent thermal spectra, *Nucl. Sci. Engng.* 10, 190–201 (1961).

97 Kaplan, S., On the best method for choosing the weighting functions in the method of weighted residuals, *Trans. Amer. Nucl. Soc.* 6, 3-4 (June 1963).

98 Clark, M. and Hansen, K. F., "Numerical methods of reactor analysis," Academic Press, 1964.

99 Roberts, P. H., Characteristic value problems posed by differential equations arising in hydrodynamics and hydromagnetics, *J. Math. Anal. Applic.* 1, 195–214 (1960).

100 DiPrima, R. C., Some variational principles for problems in hydrodynamic and hydromagnetic stability, *Quart. Appl. Math.* 18, 375–385 (1961); AMR 14(1961), Rev. 4462.

101 Glansdorff, P., Prigogine, I. and Hays, D. F., Variational properties of a viscous liquid at a nonuniform temperature, *Physics of Fluids* 5, 144–149 (1962).

102 Glansdorff, P. and Prigogine, I., On a general evolution criterion in macroscopic physics, *Physica* 30, 351–374 (1964).

103 Prigogine, I. and Glansdorff, P., Variational properties and fluctuation theory, *Physica* 31, 1242–1256 (1965).

104 Prigogine, I., Evolution criteria, variational properties, and fluctuations, pp. 3-16 in Ref. 68.

105 Schechter, R. S., On a variational principle for the Reiner-Rivlin fluid, *Chem. Engng. Sci.* 17, 803–806 (1962).

106 Hays, D. F., A variational formulation applied to Couette and Poiseuille flow, *Acad. Roy. Belg., Bull. Classe Sci.* 49, 576–602 (1963), AMR 17(1964), Rev. 4702.

107 Glansdorff, P., Dekeuster, E. and Haezendonck, Y., Application du potentiel local de production d'entropie à la recherche de distributions stationnaires de températures et de concentrations, *Acad. Roy. Belg., Bull. Classe Sci.* 50, 60–73 (1964).

108 Hays, D. F., Application of a variational formulation for time-dependent and steady-state solutions of the heat equation: temperature-dependent thermal conductivity, pp. 17–43 in Ref. 68.

109 Schechter, R. S. and Himmelblau, D. M., Local potential and system stability, *Physics of Fluids* 8, 1431–1437 (1965).

110 Rosen, P., On variational principles for irreversible processes, *J. Chem. Phys.* 21, 1220–1221 (1953).

111 Rosen, P., Use of restricted variational principles for the solution of differential equations, *J. Appl. Phys.* 25, 336–338 (1954).

112 Rosen, P., The solution of the Boltzmann equation for a shock wave using a restricted variational principle, *J. Chem. Phys.* 22, 1045–1049 (1954).

113 Rosen, P., Variational approach to magnetohydrodynamics, *Physics of Fluids* 1, 251 (1958); AMR 12(1959), Rev. 1031.

114 Finlayson, B. A. and Scriven, L. E., Galerkin's method and the local potential, pp. 291–294 in Ref. 68.

115 Biot, M. A., New methods in heat flow analysis with application to flight structures, *J. Aero. Sci.* 24, 857–873 (1957).

116 Biot, M. A., Further developments of new methods in heat-flow analysis, *J. Aero/Space Sci.* 26, 367–381 (1959); AMR 13(1960), Rev. 1376.

117 Biot, M. A., "Thermodynamics and heat-flow analysis by Lagrangian methods," in Seventh Anglo-American Aeronautical Conference, Inst. Aero. Sci. 1960, pp. 418–431.

118 Biot, M. A., Variational and Lagrangian thermodynamics of thermal convection: fundamental shortcomings of the heat-transfer coefficient, *J. Aero/Space Sci.* 29, 105–106 (1962).

119 Biot, M. A., Lagrangian thermodynamics of heat transfer in systems including fluid motion, *J. Aero/Space Sci.* 29, 568–577 (1962); AMR 16(1963), Rev. 5413.

120 Biot, M. A. and Daughaday, H., Variational analysis of ablation, *J. Aero/Space Sci.* 29, 227–229 (1962); AMR 15(1962), Rev. 5487.

121 Biot, M. A. and Agrawal, H. C., Variational analysis of ablation for variable properties, *Trans. ASME, J. Heat Transfer* 86, 437–442 (1964); AMR 18(1965), Rev. 4989.

122 Nigam, S. D. and Agrawal, H. C., A variational principle for convection of heat. I, *J. Math. Mech.* 9, 869–872 (1960); AMR 14(1961), Rev. 4982.

123 Agrawal, H. C., A variational principle for convection of heat, II, *J. Math. Mech.* 9, 872–884 (1960); AMR 14(1961), Rev. 4982.

124 Gupta, S. C., Variational principle for fully developed laminar heat transfer in uniform channels, *Appl. Sci. Res.* A10, 85–101 (1961).

125 Gupta, S. C., A variational principle for convection of heat in anisotropic media, *Appl. Sci. Res.* A10, 229–234 (1961); AMR 15(1962), Rev. 1680.

126 Gupta, S. C., A least-squares procedure for the variational principle for heat conduction in anisotropic media, *Appl. Sci. Res.* A11, 311–312 (1963); AMR 17(1964), Rev. 2865.

127 Lardner, T. J., Biot's variational principle in heat conduction, *AIAA J.* 1, 196–206 (1963).

128 Tao, L. N., On the variational principle and Lagrange equations in studies of gasdynamic lubrication, *J. Appl. Mech.* 31, 43–46 (1964); AMR 18(1965), Rev. 7173.

129 Fujita, H., Diffusion with a sharp moving boundary, *J. Chem. Phys.* 21, 700–705 (1953).

129a Finlayson, B. A. and Scriven, L. E., Upper and lower bounds for solutions to the transport equation, *AIChE J.*

129b Finlayson, B. A. and Scriven, L. E., "On the search for variational principles," to be published.

Additional References to Applications

Nonlinear Problems

Galerkin method: Refs. 13, 22, 53, 106–108, 115, 121 and

130 Fung, Y. C., On two-dimensional panel flutter, *J. Aero. Sci.* 25, 145–160 (1958).

131 Lourié, A. and Tchekmarev, A., Sur la solution approximative de quelques problèmes nonlinéaires de vibrations forcées" (in Russian with French summary), *Prikl. Mat. Mekh.* 1, p. 324 (1937).

132 Panov, D. J., On an application of Galerkin's method to certain nonlinear problems in the theory of elasticity (in Russian with English summary), *Prikl. Mat. Mekh.* 3, 139–142 (1939).

133 Richardson, P. D., Unsteady one-dimensional heat conduction with a nonlinear boundary condition, *J. Heat Transfer* 86, 298–299 (1964).

134 Schwesinger, G., On one-term approximations of forced nonharmonic vibrations, *J. Appl. Mech.* 17, 202–208 (1950).

135 Tanner, R. I., "The estimation of bearing loads using Galerkin's method," *Internat. J. Mech. Sci.* 3, 13–27 (1961); AMR 15(1962), Rev. 1222.

Integral method: Refs. 20, 34–36, and

136 Acrivos, A., Shah, M. J. and Petersen, E. E., On the solution of the two-dimensional boundary-layer flow equations for a non-Newtonian power law fluid, *Chem. Engng. Sci.* 20, 101–105 (1965).

137 Buzzell, G. D. and Slattery, J. C., Non-Newtonian boundary layer flow, *Chem. Engng. Sci.* 17, 777–782 (1962).

138 Bogue, D. C., Entrance effects and prediction of turbulence in non-Newtonian flow, *Ind. Engng. Chem.* 51, 874–878 (1959).

139 Campbell, W. D. and Slattery, J. C., Flow in the entrance of a tube, *Trans. ASME, J. Basic Engng.* 85, 41–46 (1963); AMR 17(1964), Rev. 3956.

140 Goodman, T. R., The heat-balance integral and its application to problems involving a change in phase, *Trans. ASME* 80, 335–342 (1958).

141 Goodman, T. R., "The heat-balance integral: further considerations and refinements," *Trans. ASME, J. Heat Transfer* 83, 83–86 (1961), AMR 14(1961), Rev. 4369.

142 Goodman, T. R. and Shea, J. J., The melting of finite slabs, *J. Appl. Mech.* 27, 16–24 (1960); AMR 12(1959), Rev. 1096.

143 Kapur, J. N. and Gupta, R. C., Two-dimensional flow of power-law fluids in the inlet length of a straight channel. I and II, *ZAMM* 43, 135–141 (1963), 44, 277–284 (1964); AMR 18(1965), Rev. 7550.

144 Kapur, J. N. and Srivastava, A. C., Axially symmetric and two-dimensional stagnation point flows of a certain visco-elastic fluid, *J. Phys. Soc. Japan* 18, 441–444 (1963); AMR 17(1964), Rev. 1530.

145 Lemlich, R. and Steinkamp, J. S., Laminar natural convection to an isothermal flat plate with a spatially varying acceleration, *AIChE J.* 10, 445–447 (1964); AMR 18(1965), Rev. 3021.

146 Rajeshwari, G. K. and Rathna, S. L., Flow of a particular class of non-Newtonian visco-elastic and visco-inelastic fluids near a stagnation point, *ZAMP* 13, 43–57 (1962); AMR 17(1964), Rev. 872.

147 Rosenhead, L., editor, "Laminar boundary layers," Clarendon Press, 1963, esp. Chapt. 6.

148 Schlichting, H., "Boundary layer theory," 4th edn., McGraw-Hill, 1960, esp. Chapt. 12.

149 Siegel, R. and Sparrow, E. M., Simultaneous development of velocity and temperature distributions in a flat duct with uniform wall heating, *AIChE J.* 5, 73–75 (1959); AMR 13(1960), Rev. 405.

150 Sowerby, L., An application of Pohlhausen's method to a problem of flow along an edge, *Appl. Sci. Res.* A12, 417–424 (1964), AMR 18(1965), Rev. 1647.

151 Srivastava, A. C., The flow of a non-Newtonian liquid near a stagnation point, *ZAMP* 9, 80–84 (1958).

152 Thwaites, B., Approximate calculation of the laminar boundary layer, *Aeronaut. Quart.* 1, 245–280 (1949).

153 Yang, K. T. and Szewczyk, A., An approximate treatment of unsteady heat conduction in semi-infinite solids with variable thermal properties, *Trans. ASME, J. Heat Transfer* 81, 251–252 (1959); AMR 13(1960), Rev. 6504.

Subdomain method: Ref. 54 and

154 Carter, L. F. and Gill, W. N., Asymptotic solution for combined free and forced convection in vertical and horizontal conduits with uniform suction and blowing, *AIChE J.* 10, 330–339 (1964).

155 Launder, B. E., An improved Pohlhausen-type method of calculating the two-dimensional boundary layer in a pressure gradient," *Trans. ASME, J. Heat Transfer* 86, 360–364 (1964); AMR 18(1965), Rev. 1004.

156 Dorodnitsyn, A. A., "On a method of numerical solution of some nonlinear problems in aero-hydrodynamics," Actes Ninth Intern. Congr. Appl. Mech., Brussels, 1, 485 (1957).

157 Dorodnitsyn, A. A., "A contribution to the solution of mixed problems of transonic aerodynamics," in Advances in the Aeronautical Sciences, Vol. 2 of the Proc. First Intern. Congr. Aero. Sci., Madrid, 1958, Pergamon, 1959, pp. 832–844.

Least-squares method: Refs. 50 and 134.

Method of moments: Refs. 16, 129 and

158 Poots, G., An approximate treatment of a heat conduction problem involving a two-dimensional solidification front, *Internat. J. Heat Mass Transfer* 5, 339–348 (1962); AMR 16(1963), Rev. 433.

159 Prager, S., Transport processes with chemical reactions, *Chem. Engng. Prog., Sympos. Series* 55, no. 25, 11–22 (1959).

Linear Initial-value Problems

Galerkin method: Refs. 12, 25, 39, 60, 63, 108, 115.

Integral method: Refs. 34, 48, 147 and

160 Goodman, T. R., The heating of slabs with arbitrary heat input, *J. Aero/Space Sci.* 26, 187–188 (1959); AMR 12(1959), Rev. 4095.

161 Lyman, F. A., Heat transfer at a stagnation point when the free-stream temperature is suddenly increased, *Appl. Sci. Res.* A13, 65–80 (1964); AMR 18(1965), Rev. 4300.

162 Sparrow, E. M., Unsteady stagnation-point heat transfer, NASA TN D-77 (1959); AMR 13(1960), Rev. 5374.

163 Sparrow, E. M. and Siegel, R., "Thermal entrance region of a circular tube under transient heating conditions," Proc. Third U.S. Natl. Congr. Appl. Mech. 1958, Providence, R. I., pp. 817–826; AMR 12(1959), Rev. 5207.

Least-squares method: Refs. 12, 50, 51.

Method of moments: Refs. 12, 48, 129.

Collocation: Ref. 12.

Linear Boundary-value Problems

Galerkin method: Refs. 4, 37, 40, 112, 120, 124, 127, and

164 Kaplan, S. and Marlowe, O. J., Application of synthesis approximations to three-dimensional depletion calculations and to cell theory, *Trans. Amer. Nucl. Soc.* 6, 254–255 (Nov. 1963).

165 Wernick, R. J. and Pan, C. H. T., Static and dynamic characteristics of self-acting, partial arc, gas journal-bearings, *Trans. ASME, J. Basic Engng.* 86, 405–413 (1964); AMR 17(1964), Rev. 7587.

Integral method: Refs. 147, 148 and

166 Devan, L. and Oberai, M. M., Approximate solution of second-order boundary-layer equations, *AIAA J.* 2, 1838–1840 (1964); AMR 18(1965), Rev. 2256.

167 Libby, P. A., Morduchow, M. and Bloom, M., "Critical study of integral methods in compressible laminar boundary layers," NACA TN 2655 (1952).

168 Morduchow, M., On laminar flow through a channel or tube with injection: application of the method of averages, *Quart. Appl. Math.* 14, 361–368 (1957).

169 Ruckenstein, E., On mass transfer in the continuous phase from spherical bubbles or drops, *Chem. Engng. Sci.* 19, 131–146 (1964).

Method of moments:

170 Beek, W. J. and Kramers, H., Mass transfer with change in interfacial area, *Chem. Engng. Sci.* 16, 909–921 (1962).

Collocation method: Ref. 4 and

171 Conway, H. D., The approximate analysis of certain boundary-value problems, *Trans. ASME* 82, 275–277 (1960); AMR 13(1960), Rev. 6139.

172 Jain, M. K., An extremal point collocation method for fluid flow problems, *Appl. Sci. Res.* A11, 177–188 (1962); AMR 16(1963), Rev. 4991.

173 Jeffreys, H., The surface elevation in cellular convection, *Quart. J. Mech. Appl. Math.* 4, 283–288 (1951).

Linear Eigenvalue Problems

[illegible]

[illegible], H. and Zartarian, G., Piston theory: a new aerodynamic tool for the aeroelastician, *J. Aero. Sci.* 23, 1109–1118 (1956).

175 Cheng, H. S. and Pan, C. H. T., Stability analysis of gas-lubricated, self-acting, plain, cylindrical journal bearings of finite length, using Galerkin's method, *Trans. ASME, J. Basic Engng.* 87, 185–192 (1965).

176 Cheng, H. S. and Trumpler, P. R., Stability of the high-speed journal bearing under steady load, *Trans. ASME, J. Engng. Industr.* 85, 274–280 (1963).

177 DiPrima, R. C., Application of the Galerkin method to problems in hydrodynamic stability, *Quart. Appl. Math.* 13, 55–62 (1955).

178 DiPrima, R. C., The stability of viscous flow between rotating concentric cylinders with a pressure gradient acting round the cylinders, *J. Fluid Mech.* 6, 462–468 (1959); AMR 15(1962), Rev. 283.

179 DiPrima, R. C., The stability of viscous fluid between rotating cylinders with an axial flow, *J. Fluid Mech.* 9, 621–631 (1960); AMR 14(1961), Rev. 4272.

180 DiPrima, R. C., Stability of nonrotationally symmetric disturbances for viscous flow between rotating cylinders, *Phys. Fluids* 4, 751–755 (1961); AMR 15(1962), Rev. 1530.

181 DiPrima, R. C. and Pan, C. H. T., The stability of flow between concentric cylindrical surfaces with a circular [illegible]

* 182 Duncan, W. J., Note on Galerkin's method for the treatment of problems concerning elastic bodies, *Phil. Mag.* (7) 25, 628–633 (1938).

183 Duncan, W. J., Applications of the Galerkin method to the torsion and flexure of cylinders and prisms, *Phil. Mag.* (7) 25, 634–649 (1938).

184 Duncan, W. J. and Lyon, H. M., "Torsional oscillations of a cantilever when the stiffness is of composite origin," Gt. Brit. Aero. Res. Council Rept. and Memo. 1809 (1937); reprinted in Gt. Brit. Air Ministry Aero. Res. Comm. Tech. Rept. 1, 471–483 (1937).

185 Hedgepeth, J. M., On the flutter of panels at high Mach numbers, *J. Aero. Sci.* 23, 609 (1956).

186 Hedgepeth, J. M., Flutter of rectangular simply supported panels at high supersonic velocities, *J. Aero. Sci.* 24, 563–573 (1957).

187 Rattayya, J. V., Flutter analysis of circular panels, *J. Aero/Space Sci.* 29, 534–539 (1962); AMR 15(1962), Rev. 7342.

Least-squares method: Refs. 4, 50, 64.

Collocation method: Ref. 64.

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

USOS DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPRESION EN AGUAS, COSTERAS LAGOS Y RIOS

PROBLEMAS AMBIENTALES ASOCIADOS EN FLUJO.

MAYO, 1978.

ENVIRONMENTAL PROBLEMS ASSOCIATED WITH FLUID FLOW

R. H. Gallagher and D. L. Young

I. INTRODUCTION

The topic of computational fluid mechanics does not have a long history. By the 1930's there were only a limited number of papers on this topic, a condition that was due not only to the absence of high speed computational facilities but also to difficulties in dealing with the inherent nonlinearity of most problems of interest. Thus, computational fluid mechanics has developed in tandem with the finite element method. Very many applications of the latter to flow problems have been recorded, as described in References 1-4. As these references demonstrate, a detailed review of the complete field would require a complete text. We therefore limit our attention, in these lecture notes, to the application of the finite element method to environmental problems associated with fluid flow.

A large number of conditions can be grouped under the heading of "environmental problems". In the present case we refer principally to the transport of heat or the concentration of a substance through a body of water. The processes of convection and diffusion participate in the transport process. Velocities appear in the convective terms and, although the most vigorous treatment of the problem will involve coupling of velocity and temperature (or concentration) equations, practical considerations may require independent solutions for the two fields. We therefore include in our review a discussion of solutions for flow velocities alone for lake and stream situations.

A study of the literature of topic under review discloses that although the problems are basically three-dimensional, no numerical solutions of this scale have yet been attempted. Simplifying assumptions are customarily made about one of the dimensions and the problem is reduced to one of analysis in the plane. Indeed, some investigations make assumptions regarding two of the dimensions and study the velocity in one dimension together with the temperature.

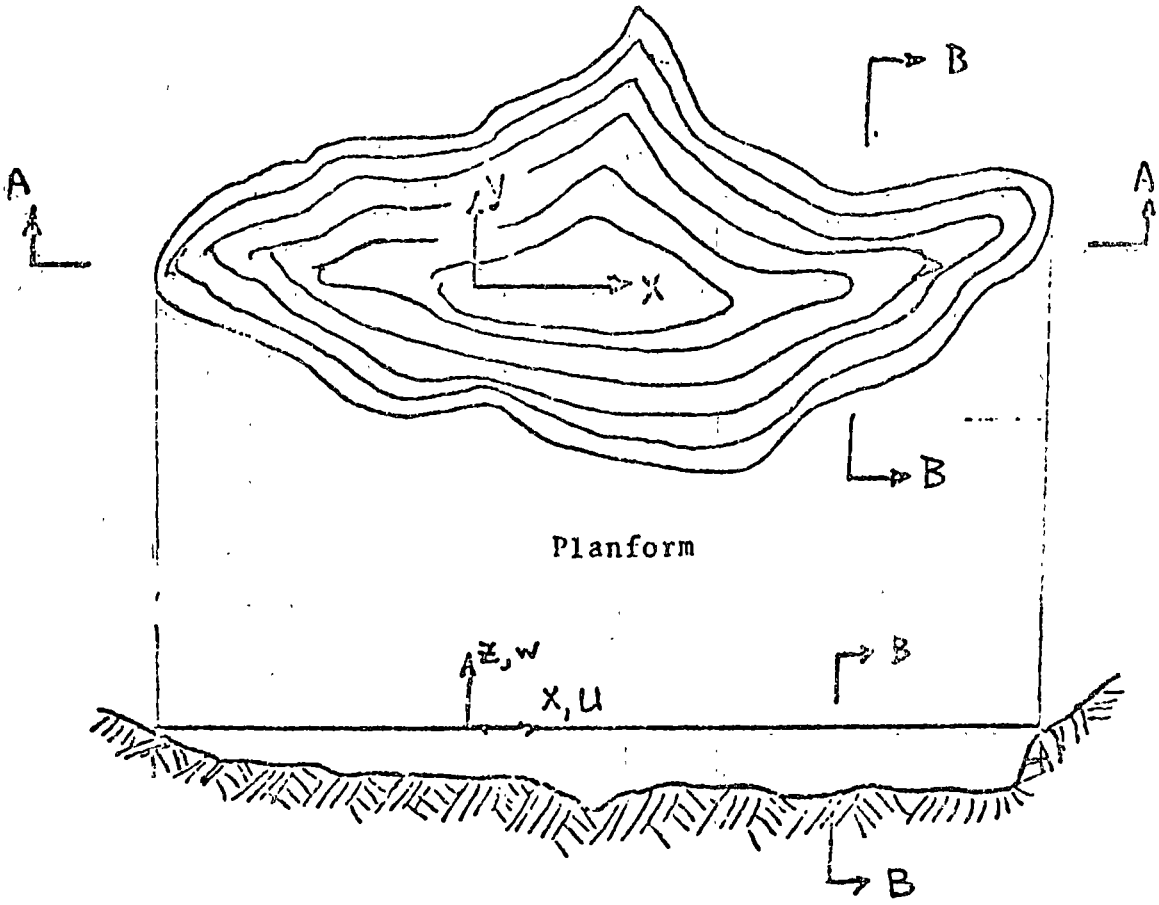
In view of the above circumstance, these notes are categorized with respect to the different types of two-dimensional situations. Only the cases of flow in planform and that of flow on the narrow cross-section of a lake or similar body of water are treated in this review.

First we define the coordinate systems associated with the respective types of problems. Then, separate sections are devoted to each type.

II. COORDINATE SYSTEMS AND GOVERNING EQUATIONS

Figure 1 illustrates the body of water and the associated coordinate systems. The body of water we have in mind is a lake, although cases will be treated which refer to streams and estuaries. In the latter circumstances the flow is predominantly in the y-direction.

The planform (x-y plane) is the basis for analysis of wind-driven circulation and of flow through basins and estuaries. The assumptions that are invoked are discussed in some detail later, but, for the present we simply note that they are directed to elimination of the z-coordinate from the problem. The maximum



Section A-A

Figure 1 Lake-Planform and Side Views

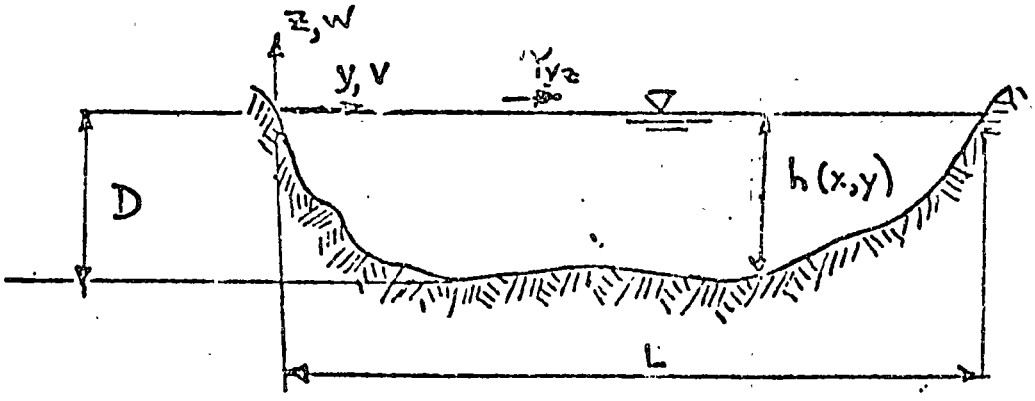


Figure 2 Narrow-Cross-Section (Section B-B)
(Vertical and Horizontal Dimensions have been Exaggerated in Comparison with Fig. 1)

z-dimension is very much smaller than the y- and x-dimensions. The latter may be of approximately the same magnitude.

The narrow cross-section (x-z) is intended to represent the section of a lake or similar body of water. Here, the x- and z-dimensions are of similar magnitude and the y-direction is very large. Finally, we have the side view (y-z), which refers principally to flow in streams and estuaries and which is often reduced to just the y-direction.

Physically, the velocity and temperature fields are determined through the conservation of mass, momentum and energy. The governing equations are

Continuity (Conservation of mass)

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \qquad (1)$$

x-Momentum

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} - fu = \frac{1}{\rho_o}\frac{\partial p}{\partial x} + \frac{\partial}{\partial x}\left(K^M_{xx}\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(K^M_{xy}\frac{\partial u}{\partial y}\right) + \frac{\partial}{\partial z}\left(K^M_{xz}\frac{\partial u}{\partial z}\right) \qquad (2)$$

and correspondingly for y and z. (In the z-direction there is no Coriolis force (- fu) but a buoyancy term $(\frac{\rho}{\rho_o} g_x)$ must be added to the left side.)

Temperature (Conservation of Energy)

$$\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} + w\frac{\partial T}{\partial z} = \frac{\partial}{\partial x}\left(K^H_x\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(K^H_y\frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(K^H_z\frac{\partial T}{\partial z}\right) \qquad (3)$$

In these equations u, v, and w are the x, y, and z-direction velocities, g is the gravitational acceleration, p is pressure, ρ is density, ρ_o is reference density, T is temperature. f is the Coriolis parameter and K^M and K^H (with appropriate directional

subscripts) are the viscosity and diffusivity. The above equations are supplemented by equations of state, such as density as a function of temperature, pressure and concentration of a substance, and the viscosities and diffusivities as a function of "stability parameters", e.g., Richardson Number, Prandtl Number, Monin-Obukhov Length, etc.

These are very general equations. Consequently, in fact, the following approximations may be introduced, depending on the analyst's interests and goal.

1) The Boussinesq approximation: The variation of density is small, so that the fluid can be treated as an incompressible fluid. The variation of density is only considered in the buoyancy term $\frac{\rho}{\rho_o} g$ in the z-momentum equation.

2) The shallow water approximation: The inertia forces are negligible compared to the other forces. Also, the w-component is much smaller than the horizontal components, so that the pressure is hydrostatic ($g = -\frac{1}{\rho}\frac{\partial p}{\partial z}$, where g is the acceleration due to gravity).

3) The eddy viscosity and diffusivity approximations: Since the stratification is almost perpendicular to the gravitational force, it is customary to assume that the horizontal eddy viscosity (K^M_{xx}, etc.) and diffusivity (K^H_x) are approximated by constants, while the vertical ones are functions of the gradients of density and velocity. The exact relationships are still hot debate. In practical analysis, the determination must come from semi-empirical stratified turbulent theory. (Monin-Yaglom, Ref. 5)

III. PLANFORM ANALYSIS. WIND-DRIVEN CIRCULATION AND FLOW THROUGH BASINS

The cross-section shown in Figure 2 defines the basic geometric parameters of this development, which is due to Liggett and Hadjitheodourou[6] in its fundamental theoretical form. The origin of coordinates is fixed at the surface of the lake, with velocity w = 0 at z = o. (The 'rigid lid' assumption). The physical properties of the lake, including the eddy viscosity and the mass density per unit volume, are assumed to be constant and the Coriolis parameter is also assumed constant. The pressure is taken to vary hydrostatically. The surface wind stresses τ_{xz} and τ_{yz} are prescribed. Under these assumptions the momentum equations take the form

$$-fv = -\frac{1}{\rho_o}\frac{\partial p}{\partial x} + \frac{\partial^2 u}{\partial z^2} \cdot K_o^M \tag{4}$$

$$fu = -\frac{1}{\rho_o}\frac{\partial p}{\partial y} + \frac{\partial^2 v}{\partial z^2} \cdot K_o^M \tag{5}$$

$$g = -\frac{1}{\rho_o}\frac{\partial p}{\partial z} \tag{6}$$

The continuity equation is unmodified.

A stream function ψ is defined as follows

$$\frac{\partial \psi}{\partial y} = \bar{u}h \tag{7}$$

$$-\frac{\partial \psi}{\partial x} = \bar{v}h \tag{8}$$

in which $\bar{u}$ and $\bar{v}$ are depthwise averages of the component velocities. After combination of the above equations, with consideration of the boundary conditions (zero velocity on all solid surfaces and $\tau_{xz} = K_o^M \frac{\partial u}{\partial x}$, $\tau_{yz} = K_o^M \frac{\partial v}{\partial y}$) one obtains

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + A(x,y)\,\frac{\partial\psi}{\partial x} + B(x,y)\,\frac{\partial\psi}{\partial y} + C(x,y) = 0 \qquad (9)$$

The terms $A(x,y)$, $B(x,y)$ and $C(x,y)$ are those which result from consideration of the varying depth and, as indicated, are functions of the planform coordinates x and y. Thus, they account for the varying depth.

Transformation of Eq. (9) into the finite element form, given in detail in Ref. 7, is accomplished by means of the Galerkin method. The approximation of ψ is by means of the trial function $\bar{\psi}$, which has the form

$$\bar{\psi} = N_i\psi_i = \lfloor N\rfloor\{\psi\} \qquad (10)$$

wherein the N_i are the shape functions and $\{\psi\}$ are the nodal values of the stream function. Applying now the weighted residual concept

$$\int_A \{N\}\left(\frac{\partial^2\lfloor N\rfloor}{\partial x^2} + \frac{\partial^2\lfloor N\rfloor}{\partial y^2} + \frac{A\partial\lfloor N\rfloor}{\partial x} + \frac{B\partial\lfloor N\rfloor}{\partial y} + C\right)dA\ \{\psi\} = 0 \qquad (11)$$

Next, integration by parts in the plane is applied to reduce the order of the derivatives appearing in this integral and to produce boundary terms. One obtains

$$\int_A \left[-\frac{\partial}{\partial x}\{N\}\,\frac{\partial\lfloor N\rfloor}{\partial x} - \frac{\partial}{\partial y}\{N\}\,\frac{\partial\lfloor N\rfloor}{\partial y} + A\{N\}\,\frac{\partial\lfloor N\rfloor}{\partial x} + B\{N\}\,\frac{\partial\lfloor N\rfloor}{\partial y}\right]\{\psi\} + \{N\}C\,dA + \oint \{N\}\,\frac{\partial\lfloor N\rfloor}{\partial n}\{\psi\}\,dS = 0 \qquad (12)$$

The values of $\{\psi\}$ are zero on the entire exterior boundary and the closure integrals along interelement boundaries vanish if continuity is preserved across these boundaries. Thus, the contour integral term is excluded from subsequent consideration.

Evaluation of the remaining integrals for all i then yields the following system of element equations

$$[k^e]\{\psi\} = \{r^e\} \tag{13}$$

in which

$$[k^e] = \int_A \left(- \frac{\partial\{N\}}{\partial x}\frac{\partial_i N_j}{\partial x} - \frac{\partial\{N\}}{\partial y}\frac{\partial_i N_j}{\partial y} + A\{N\}\frac{\partial_i N_j}{\partial x} + B\{N\}\frac{\partial_i N_j}{\partial y}\right) dA$$

$$\{r^e\} = -\int \{N\}\, C\, dA \tag{14}$$

It should be noted that due to the terms $A\{N\}\frac{\partial_i N_j}{\partial x}$ and $B\{N\}\frac{\partial_i N_j}{\partial y}$ the resulting algebraic equations will be nonsymmetric.

The equations for the complete lake are constructed from the equations of the elements by imposing the condition of stream function continuity at each element joint. Thus, the global equations are, by simple addition of all coefficients with like subscripts

$$[K]\{\psi\} = \{R\} \tag{15}$$

After solution for $\{\psi\}$ the other variables, such as averaged velocities and pressure gradients, can subsequently be evaluated by back substitution.

Numerical solutions to Eq. (15) have been obtained for both simple test problems and for Lake Ontario[7]. Since field data is not available for Lake Ontario the convergence of the solution has been studied with use of higher-order elements[8]. Cheng[9] has analyzed Lake Erie, using a formulation which excludes consideration of variable depth. Tong[10] includes this factor in a finite element formulation based on Welander's theory,[11] which does not differ significantly from the theory cited above.

If a stream function is adopted as the dependent variable, as is done in the formulations discussed previously, the presence of islands raises a basic complication in the definition of the boundary conditions at the node points of the island shore line. The stream function is zero at points on the shore of the lake but takes on a constant, undefined value on each of the islands. Thus, as Tong[10] proposes, the values of the stream function on a given island are set equal to a single value that is determined in the solution process. This substantially contracts the number of unknowns in the equations to be solved.

Cheng[9] adopts a different approach to the treatment of islands. The system of global equations is first assembled without consideration of the islands and their boundary conditions. We denote this solution as $\{\Psi_o\}$. Then, in succession, 'unit' solutions $\{\Psi_i\}$ (i=1, .. M, where M is the number of islands) are obtained for $\Psi_j = 1$ for node points on the respective islands. Finally, an M x M system of equations must be solved to give the amplitudes G_i (i=1..M) which apply to the unit solutions. The complete solution is then represented by

$$\{\Psi\} = \{\Psi_o\} + \sum_{i=1}^{M} G_i\{\Psi_i\} \qquad (16)$$

The determination of the planform distribution by transport of temperature in a lake or basin with known flow is also a problem of major practical importance, especially for cooling ponds and similar basins. Temperature distributions have been determined for such conditions by Loziuk, Anderson, and Belytschko[12,13]. Tong[10] presented a more general development

along these lines which permits the finite element calculation of any concentration of substance in a lake. We outline the latter in this section.

If we define ϕ as the average concentration across the depth (h) of the substance under study, the governing differential equation can be written as

$$\rho\left(\frac{\partial(h\phi)}{\partial t} + \frac{\bar{u}}{h}\frac{\partial(h\phi)}{\partial x} + \frac{\bar{v}}{h}\frac{\partial(h\phi)}{\partial y}\right) = \frac{\partial}{\partial x} K_x^H \frac{\partial(h\phi)}{\partial x} + \frac{\partial}{\partial y} K_y^H \frac{\partial(h\phi)}{\partial y} + Q \tag{17}$$

where K_x^H and K_y^H are the eddy diffusivity coefficients and Q is a source or sink term. Now, the approximation of ϕ can be written in the form of the trial function

$$\bar{\phi} = N_i \phi_i = \lfloor N \rfloor \{\phi\} \tag{18}$$

where $\{\phi\}$ represents nodal values of $h\phi$ and $\lfloor N \rfloor$ is the relevant set of shape functions. When the analysis is performed for temperature, with a single temperature across the depth of the lake, $T = h\phi$.

Application of the Galerkin approach can again be made to construct element equations. Using Eqs. (17) and (18), one obtains

$$[h]\{\dot{\phi}\} + [s]\{\phi\} = \{Q\} \tag{19}$$

where

$$[h] = \left[\int_A \{N\} \lfloor N \rfloor \, dA\right] \tag{20}$$

$$[s] = \left[\int_A \left[\{N\}\left(\frac{\bar{u}}{h}\frac{\partial \lfloor N \rfloor}{\partial x} + \frac{\bar{v}}{h}\frac{\partial \lfloor N \rfloor}{\partial y}\right) + K_x^H \frac{\partial\{N\}}{\partial x}\frac{\partial \lfloor N \rfloor}{\partial x} + K_y^H \frac{\partial\{N\}}{\partial y}\frac{\partial \lfloor N \rfloor}{\partial y}\right] dA\right] \tag{21}$$

The vector {Q} accounts for the source or sink terms and any prescribed boundary conditions. Finally, by assembly of the global equations from the element equations

$$[H]\{\dot{\phi}\} + [S]\{\phi\} = \{\hat{Q}\} \tag{22}$$

where [H], [S] and $\{\hat{Q}\}$ correspond to [h], [s] and {Q}.

The idealization for transport analysis is done in the same way as for flow analysis. After calculation of the velocities in the flow analysis the values obtained are used in the formation of the matrix [s].

Loziuk, et al[12,13] apply the above approach to various practical problems, including an actual lake with irregular boundary. Available field data indicate a reasonable level of agreement with the analysis results. Tong[10] calculates the diffusion of a substance in a rectangular basin containing a circular island.

Solutions for transient flow governed by the shallow water equations have been given by Connor and Wang[24]. By integrating across the depth and assuming uniform velocity and hydrostatic pressure over the depth they establish equations in terms of nodal values of flux and elevation. Solutions are given for harmonic forcing of a rectangular basin and for tidal circulation in Massachusetts Bay.

Taylor and Davis[26] have developed finite element representations of tidal propagation in estuaries. The unknowns in these equations are the node point velocities and elevations. Surface runoff, described by means of the shallow-water equations, has been studied by Al-Mashidani and Taylor (Ref. 30).

They treat a one-dimensional case, with velocity and surface elevation as problem unknowns.

Taylor and Davis (26) and Adey and Brebbia (27) have studied dispersion in estuaries. Ref. 27 uses known values of velocity and solves for the concentration. Taylor and Davis, on the other hand, solve for concentration, velocity, and surface elevation.

Planform (x-y) finite element analysis of a rather different environmental problem has been persented by Mercer and Pinder (29). They examine heat transport in the liquid and solid phases in a ground-water flow system. The finite element equations to be solved consist of two sets, one being a flow equation in terms of pressure and the second being a temperature equation. The solutions are marched in time.

IV. CROSS-SECTION ANALYSIS

The motivation for cross-section analysis (x-z) has principally been the prediction of thermal stratification, although attempts have also been made to deal with more basic phenomena in viscous flow.

Thermal stratification is widely believed to exert an important influence on lake flow phenomena through its effects on density variations and other physical factors. In many lakes uniform temperature conditions are realized in winter and, as summer atmospheric conditions approach, a rise in temperature occurs in the upper regions of the lake. The peak is reached in these regions towards the end of summer. Since the rise in

temperature penetrates to only a limited depth (say 20 to 40 feet) the lower portions of the lake are not affected, and a somewhat 'stratified' temperature profile prevails. The heated upper region is known as the epilimnion while the unheated lower region is termed the hypolimnion.

The problem to be solved is the vertical temperature profile. There is an influence, however, of the action of the wind and this produces a two-dimensional problem.

Liggett and Bedford (Ref. 14) and Bedford (Ref. 15) have dealt with the steady-state problem of a two-dimensional cavity containing a nonhomogeneous fluid subjected to surface shear. No consideration was given to eddy viscosity and diffusivity variations. The latter was accounted for by Young, Liggett, and Gallagher (Ref. 16) and the results demonstrate that stratification, as well as circulation patterns, can be predicted with the proper empirical definition of these variations. Skiba, Unny and Weaver (Ref. 17), Debongnie (Ref. 18), and Kawahara, et al (Ref. 19) have studied cavity flow without the consideration of temperature. Coupled velocity-temperature solutions are also described by Zienkiewicz, Gallagher, and Hood (Ref. 20). In the following we describe the development of Young, Liggett and Gallagher.[16]

The physical properties which enter into the differential equations of the problem are the eddy viscosity and the eddy diffusivity. The eddy viscosity and diffusivity in the horizontal direction (K^M_{xx} etc. and K^H_x) can realistically be taken as constant. Values of these coefficients are customarily taken as the same

magnitude as those which are measured under neutral stratification. The vertical eddy viscosity and diffusivity (K^M_{xz} etc. and K^H_z) vary highly within the whole basin, however, and are dependent on such factors as the turbulence level in the surface layer, the depth, the local density gradient and the overall motion with respect to the specified geometry.

No satisfactory theory for the prediction of these variations from the more basic environmental and physical parameters is presently available and dependence must be placed on empirical relationships. In this work the relationships employed are extended forms of those proposed by Sundaram and Rehm[21], as follows

$$K^M_{xz} = K^M_{zz} = K^M_o (1 - \sigma_m Ri) \tag{23}$$

$$K^H_z = K^H_o (1 - \sigma_h Ri) \tag{24}$$

where Ri, the Richardson number, is

$$Ri = - \frac{gz^2 \frac{\partial \rho}{\partial z}}{\rho_o U^2} \tag{25}$$

in which U is a characteristic velocity, σ_m and σ_h are empirical constants, and K^M_o and K^H_o are the vertical eddy viscosity and diffusivity under neutral stratification. The continuity equation, with the assumption of incompressibility, simplifies to

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 \tag{26}$$

In defining the relevant forms of the momentum equations we assume that Boussinesq's approximation applies (ρ is taken as constant except when multiplied by g, i.e., in buoyancy terms).

Thus,

$$u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial z} = - \frac{1}{\rho_o} \frac{\partial p}{\partial x} + \frac{\partial}{\partial x} \left(K_o^M \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial z} \left(K_z^M \frac{\partial u}{\partial z}\right) \qquad (27)$$

$$u \frac{\partial w}{\partial x} + w \frac{\partial w}{\partial z} = - \frac{1}{\rho_o} \frac{\partial p}{\partial z} + \frac{\partial}{\partial x} \left(K_o^M \frac{\partial w}{\partial x}\right) + \frac{\partial}{\partial z} \left(K_z^M \frac{\partial w}{\partial z}\right) - \frac{\rho}{\rho_o} g \qquad (28)$$

where p is the local pressure.

The diffusion-advection of temperature is given by

$$u \frac{\partial T}{\partial x} + w \frac{\partial T}{\partial z} = \frac{\partial}{\partial x} \left(K_o^H \frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial z} \left(K_z^H \frac{\partial T}{\partial z}\right) \qquad (29)$$

Finally, the equation of state can be written more explicitly as

$$\rho = \rho_0 [1 - \beta(T - T_o)] \qquad (30)$$

in which β is the coefficient of volumetric expansion (assumed constant) and T_o is the point about which the true relationship is linearized.

We introduce the stream function ψ in place of u and v, such that

$$u = \frac{\partial \psi}{\partial z} \quad , \quad w = - \frac{\partial \psi}{\partial x} \qquad (31)$$

The resulting two differential equations, which replace Eqs. (27-30), can then be written in terms of nondimensional variables as follows

$$D_1(\psi,\rho) = - \nabla^4 \psi + Re \frac{\partial(\nabla^2\psi, \psi)}{\partial(x,z)} - Re\ Ri_o \frac{\partial \rho}{\partial x}$$
$$+ \sigma_m\ Ri_o \left\{ \frac{\partial^2}{\partial z^2} \left(z^2 \frac{\partial \rho}{\partial z} \frac{\partial^2 \psi}{\partial z^2}\right)\right.$$
$$\left. + \frac{\partial^2}{\partial x \partial z} \left(z^2 \frac{\partial \rho}{\partial z} \frac{\partial^2 \psi}{\partial x \partial z}\right)\right\} = 0 \qquad (32)$$

and,

$$D_2(\psi,\rho) = - \nabla^2 \rho + Re\ Pr \frac{\partial(\rho,\psi)}{\partial(x,z)}$$
$$+ \sigma_h\ Ri_o \frac{\partial}{\partial z} \left\{ \left(z \frac{\partial \rho}{\partial z}\right)^2 \right\} = 0 \qquad (33)$$

where $Re = UH/K_o^M$ is the Reynolds number, $Pr = K_o^M/K_o^H$ is the turbulent Prandtl number under neutral stratification, and $Ri_o = -\Delta\rho gH/\rho_o U^2$ is the overall Richardson number. H is the depth of the cavity. All parameters and variables have been nondimensionalized, e.g., x and z have been divided by H.

To transform the above into a finite element representation we adopt shape function approximations for ψ and ρ and use the Galerkin method. Thus, with $\psi = N_i\psi_i$ and $\rho = Q_i\rho_i$ we have the following weighted integrals

$$\int_A N_i[D_1(\bar{\psi},\bar{\rho})]dA = 0 \tag{34}$$

$$\int_A Q_i[D_2(\bar{\psi},\bar{\rho})]dA = 0 \tag{35}$$

This leads to the following set of nonlinear algebraic equations

$$S^1_{ij}\psi_j + Re\ S^2_{ijk}\psi_j\psi_k + Re\ Ri_o\ S^3_{ij}\rho_j - \sigma_m\ Ri_o\ S^4_{ijk}\psi_j\rho_k + P_i = 0 \tag{36}$$

$$S^5_{ij}\rho_j + Re\ Pr\ S^6_{ijk}\psi_j\rho_k - \sigma_h\ Ri_o\ S^7_{ijk}\rho_j\rho_k = 0 \tag{37}$$

The multipliers Re, Ri_o, Pr, σ_m and σ_h have been preserved in these representations to enable identification of the source of each term.

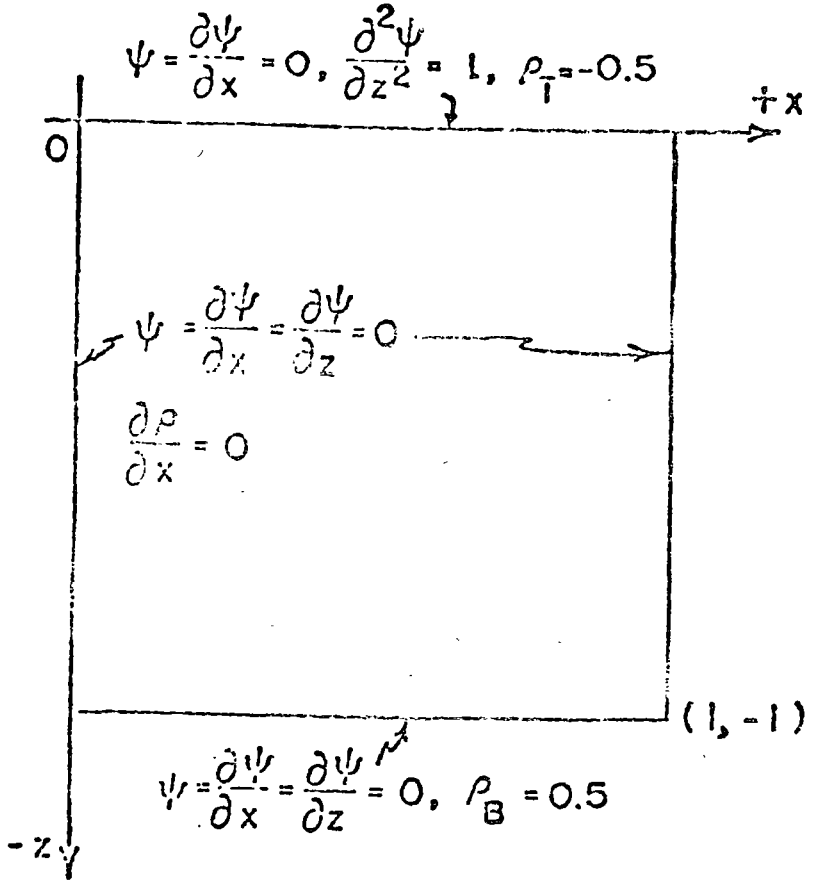
The specific algebraic form of the coefficients $S^1_{ij},\ldots,S^7_{ijk}$ is obtained after performance of the integration indicated in Eqs. (34) and (35).

The global representation is obtained by summation, from the coefficients of the above element equations, of all coefficients with like subscripts. The resulting equations are of

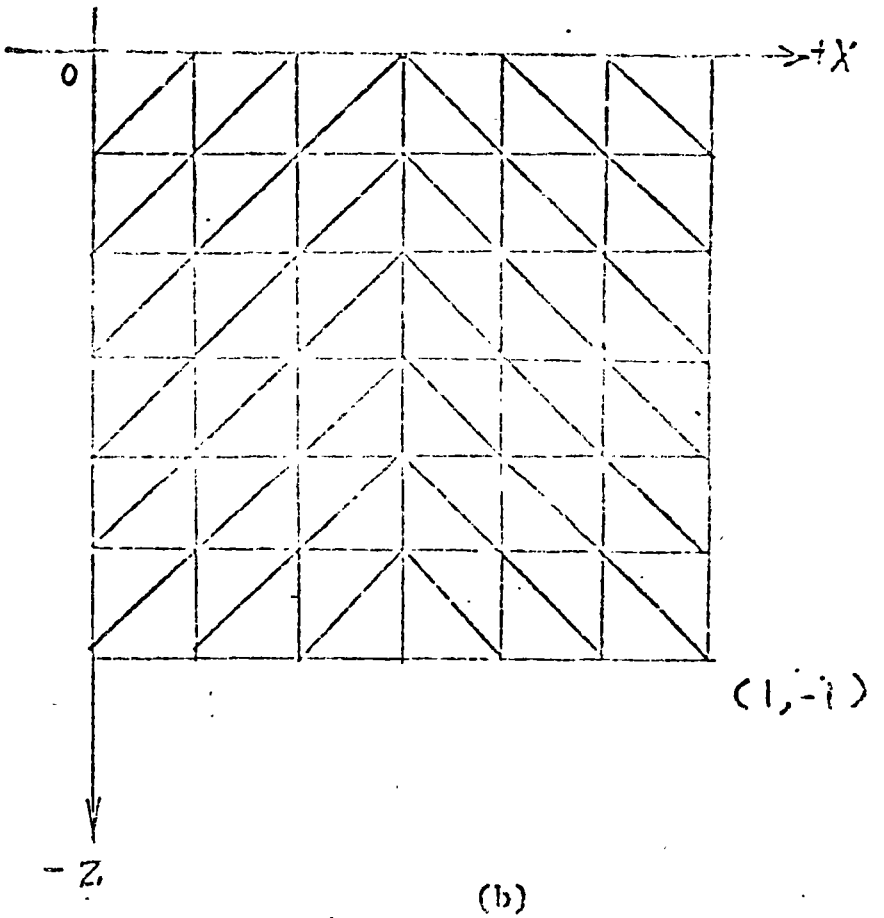
a form identical to that of Eqs. (36) and (37). The Newton-Raphson approach is adopted as the method of solution of these coupled nonlinear equations.

Numerical calculations were performed for the square cavity of Fig. 3a for the boundary conditions shown and for various assumed vertical formulations of the eddy viscosity and diffusivity. The finite element representation consisted of 72 elements arrayed in the 6x6 gridwork. (Fig. 3b).

Steady state calculations have been performed for Re = 1 to Re = 1000, Gr = 0 to Gr = 10000, and Pr = 1 to Pr = 10 where Gr is the Grashof number (The Grashof number is $Gr = Re^2 Ri_o$.). Additional numerical experiments were performed to test the sensitivity of the solution on the assumed behavior of the eddy viscosity and eddy diffusivity. Ten such computations were performed, all using Re = 100, Pr = 1, $Ri_o = 1$ but different choices of σ_h and σ_m and also different assumptions as to the form of the depthwise variation of K_{xz}^M and K_z^H as summarized in Table 1. The values of the characteristic numbers represent, of course, an infinite variety of physical data, but the following are typical: $\rho_B = 1.0$ gr/cm^3, $\rho_T = 0.9999$ gr/cm^3 ($T_T - T_B \simeq 4°C$), $\tau_x = 1.0$ dyne/cm^2, $K_o^M = K_o^H = 100$ cm^2/sec, H = 10m, and g = 980 cm/sec^2. These are approximately equivalent to the experimental data of Sundaram et al[22]. However, in the present case the boundary conditions have been chosen so that a steady-state solution exists, a condition relaxed in some subsequent computations.


$\psi = \frac{\partial \psi}{\partial x} = 0, \frac{\partial^2 \psi}{\partial z^2} = 1, \rho_T = -0.5$
0
+x
$\psi = \frac{\partial \psi}{\partial x} = \frac{\partial \psi}{\partial z} = 0$
$\frac{\partial \rho}{\partial x} = 0$
(1, -1)
$\psi = \frac{\partial \psi}{\partial x} = \frac{\partial \psi}{\partial z} = 0, \rho_B = 0.5$
-z

(a)


0
+x
(1, -1)
-z

(b)

Figure 3

TABLE 1

Summary of Computations
for Re = 100, Pr = 1, Ri_o = 1

RUN NO.	1	2	3	4	5
	σ_h	σ_m	Depth-Dependent	Cut-off Below the First Cut-off Depth	Plot Symbols
1	0	0	---	---	○
2	0.1	0.035	No	---	△
3	0.2	0.07	Yes	---	
4	0.3	0.1	Yes	---	
5	0.5	0.15	Yes	---	⊡
6	0.8	0.25	Yes	---	
7	1.0	0.35	No	No	▲
8	1.0	0.35	Yes	No	□
9	1.0	0.35	Yes	Yes	◇
10	2.0	0.70	Yes	No	●

Conclusions resulting from the first set of steady-state runs are shown in Table 2 (Ref. 15). In these runs eddy viscosity and eddy diffusivity were held constant. A typical picture of streamlines and isopycnals is shown in Figure 4. The influence of the stratification on the circulation is obvious. Additional runs could probably have elicited a specific relationship between the formation of multiple, closed circulation cells and the three parameters, Re, Pr, and Gr. However, such a relationship was not pursued since it would undoubtedly be altered with different geometries and since the eddy viscosity and eddy diffusivity relationships probably have a large effect.

That effect has been tested in ten subsequent runs which are summarized in Figures 5 and 6. In these cases the same sort of cell structure formed as shown in Figure 4, but with considerable variation in the details of the velocity, shape and size of the cells, and the density distribution.

The latter computations show that the density structure continues to have a large effect on the velocity structure and also the velocity structure greatly alters the density distribution. With the eddy viscosity and eddy diffusivity formulation that Sundaram and Rehm[21] found necessary in their one-dimensional analysis, the surface shear alone is sufficient to form a thermocline type of structure. This result is quite different, but does not conflict with, those of previous investigators who have used a one-dimensional analysis. In those previous investigations the thermocline structure formed over a period of time (several weeks) while unsteady heat inputs were applied. We have shown, however, that given an initial inhomogenicty in density, a wind shear is quite sufficient

Table 2. Summary of Results

	Streamfunction Results	Density Field Results
Re	An increasing Re reflects an increase in wind shear, the primary source of kinetic energy. Momentum transport is increasingly by convection, which discourages the formation of cells. The primary vortex center moves downstream and toward the surface with decreasing momentum diffusion.	Density transport is increasingly by advection. High Re encourages the accumulation of a thickening region of light homogeneous fluid at the top of the downwind portion of the cavity. Vertical and horizontal density gradients rise with the loss of diffusion.
Gr	At all but the lowest levels of kinetic energy (Re = 1.0) an increasing Gr encourages the growth of cells, by increasing the strength of the applied vertical density gradients. The angle of tilt between the cells diminishes with increasing Gr.	An increasing applied density difference or buoyancy strength seeks to create a linearly varying vertical density profile. Additional flow field energy is required for the readjustment as reflected in the increasing Gr.
Pr	By severely increasing the density gradients, an increased Pr encourages the formation of cells. The angle of tilt between cells increases with decreasing density diffusion.	By further diminishing density diffusion with increasing Pr an additional accumulation of lighter homogeneous fluid occurs. Also an additional increase in the energy absorbing vertical and horizontal density gradients is noted.
Low Aspect Ratio (shallow cavity)	The unimportance of horizontal diffusion results in a more exaggerated density and flow field response to any change in the Re, Pr, and Gr. Therefore cells form at lower Gr, or Pr. Vortex centers are convected farther apart horizontally with increasing transport by inertia. For a given wind and applied density difference, the vertical density gradient is more severe and thereby requires more kinetic energy to overcome.	
High Aspect Ratio (deep cavity)	Homogeneous circulation cells appear when momentum diffusion can't penetrate the wall shear resistance. Each cell rotates in the opposite direction to and is 0(10) less intense than the cell above.	All test cases were homogeneous.

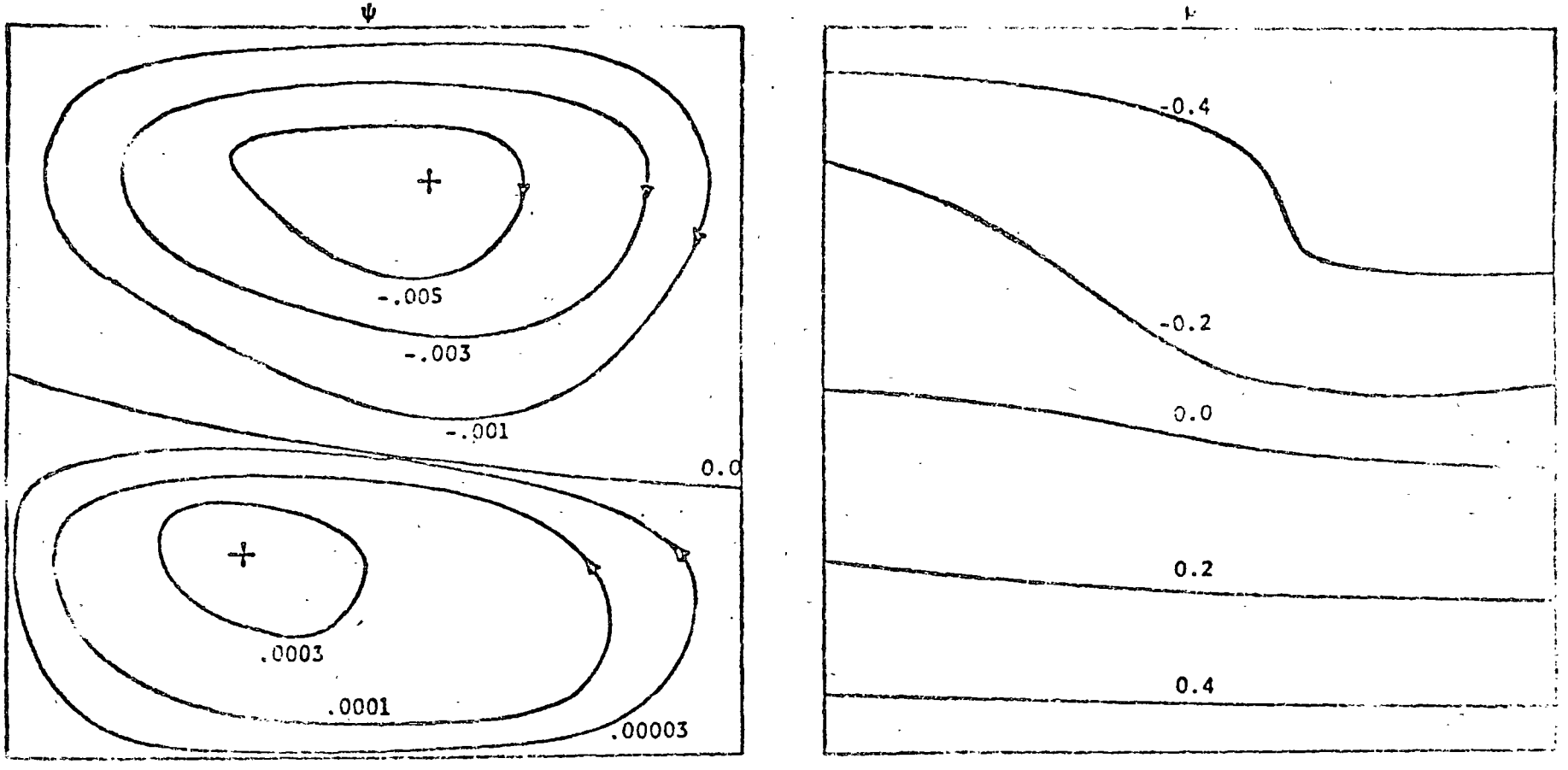


Figure 4. Streamline and Isopycnic Contours; Shear Driven, O(1) Cavity, Re = 100, Gr = 1000, Pr = 5.0

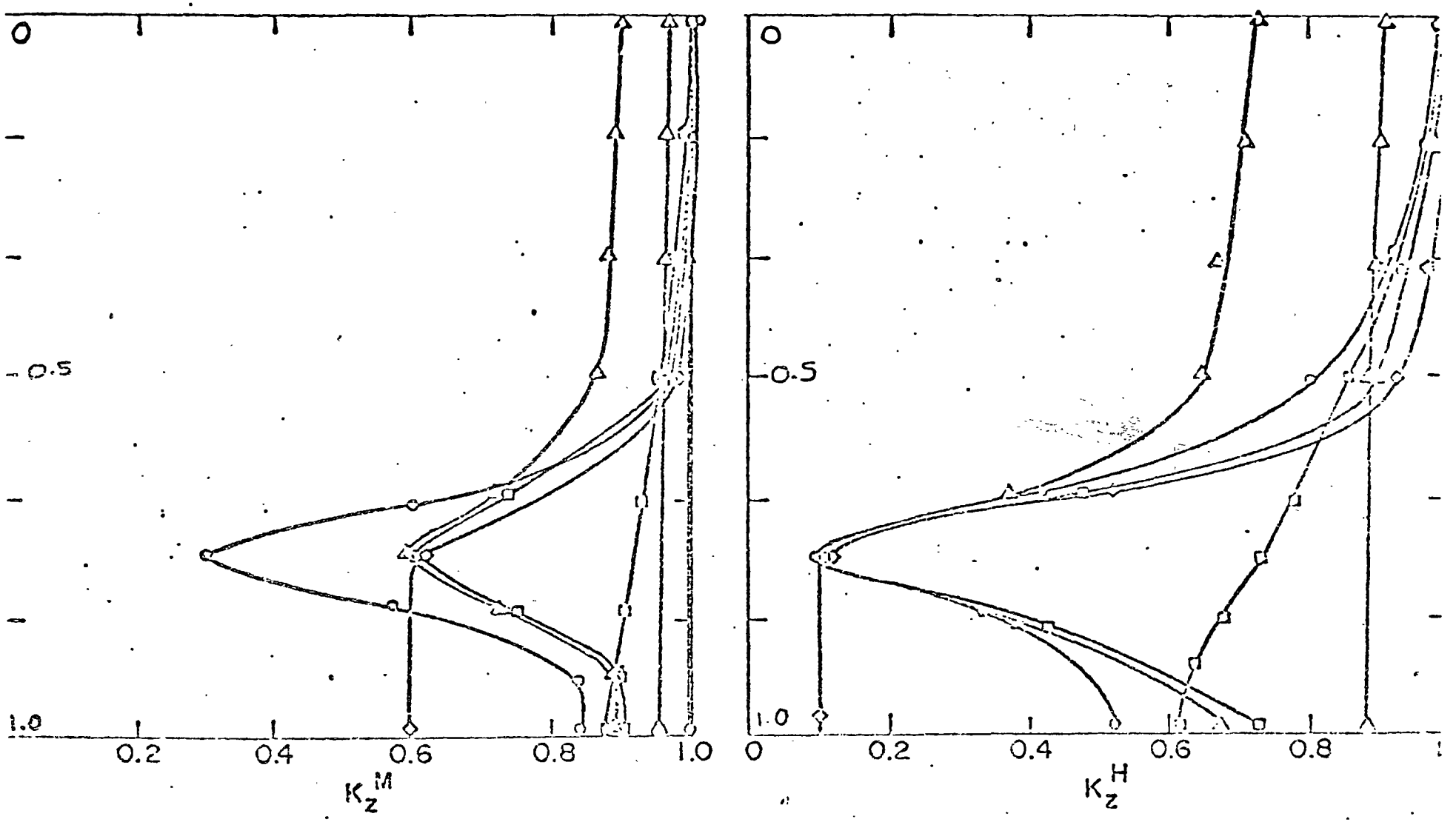
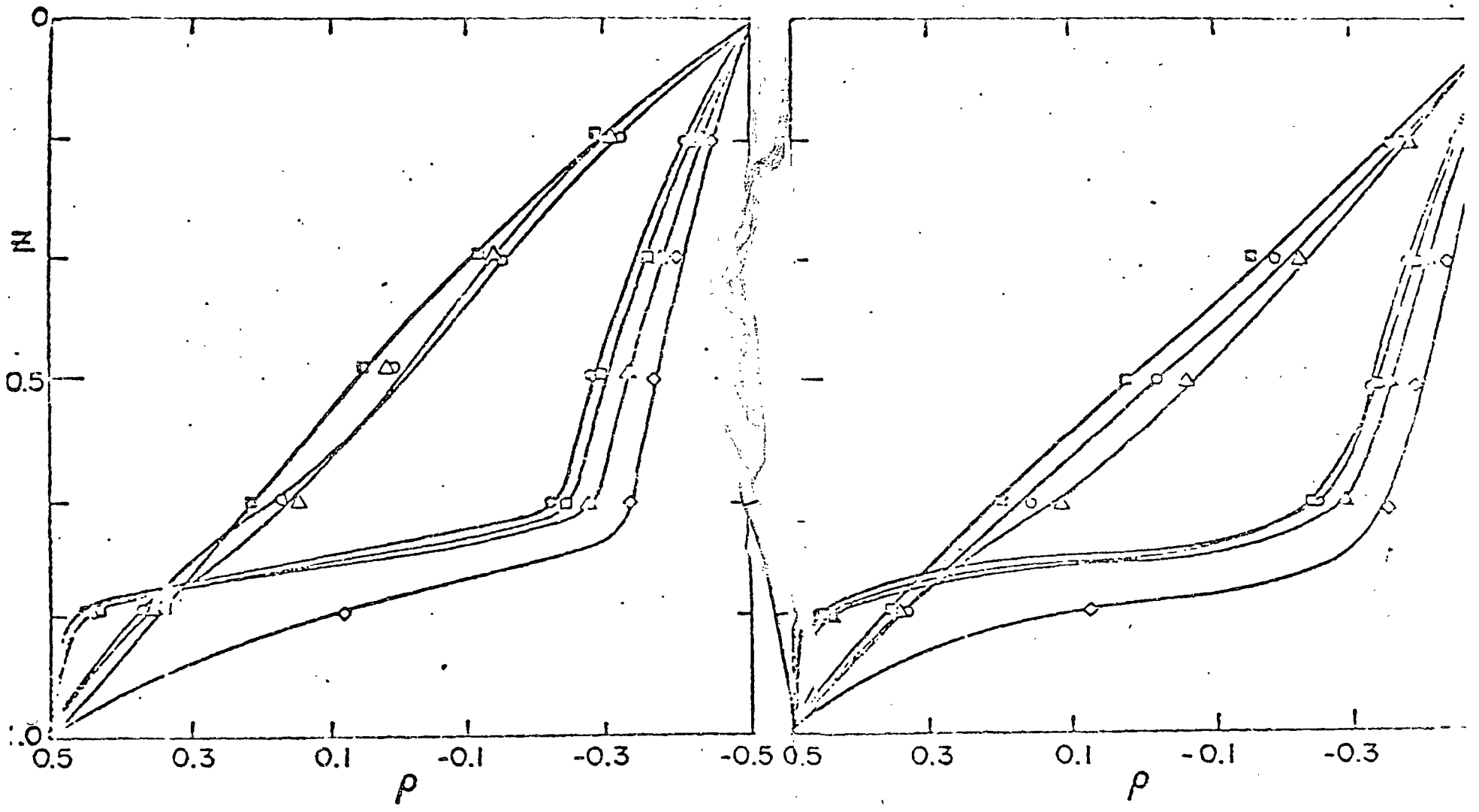


Figure 5 Vertical Distribution of Eddy Viscosity (K_z^M) and Eddy Diffusivity (K_z^H) at x = 0.8333.

(See Table 1 for definition of symbols)



a. x = 0.1667

b. x = 0.8333

Figure 6 Density (ρ) versus Depth (z) at x = 0.1667

(See Table 1 for definition of symbols)

to form the thermocline. Unsteady computations, shown below, indicate the time scale involved in such a formation and also the extent of the feedback which influences the current structure.

Fewer transient computations of cavity flow were made due to the computer costs. A total of five runs with Reynolds numbers of 100 and 1000 and Richardson numbers of 1 and 10 were made. Two values of σ_h were used. Figure 7 indicates the results of one of these calculations. In all cases in which the motion begins from rest, the entire cavity begins to circulate as a whole; that is, the cavity forms a single circulation cell. As time progresses the flow may break up into two or more cells, as is indicated in Figure 7. At the same time the density distribution is altered to show the typical thermocline shape.

The flow does not change from a state of rest to the final cell formation montonically. Instead the velocities increase rapidly to a value not far from the steady state value and then oscillate about this value. The frequency of oscillation is near the Brunt-Vaisala frequency. Other characteristics of the flow, the density gradient, the cell location, and the streamline positions, show similar damped oscillations.

The number of cells can be calculated, using certain gross approximations, from the theory of Turner[32] as expanded for this problem by Young[31]. This theory has been compared with the transient and steady-state computations with rough agreement. The difficulty in the application of such theories to real lakes (or even cavities) is that all the factors, the most important being the density distribution, and the interaction of those factors cannot te considered adequately. Results indicate that multiple cells are

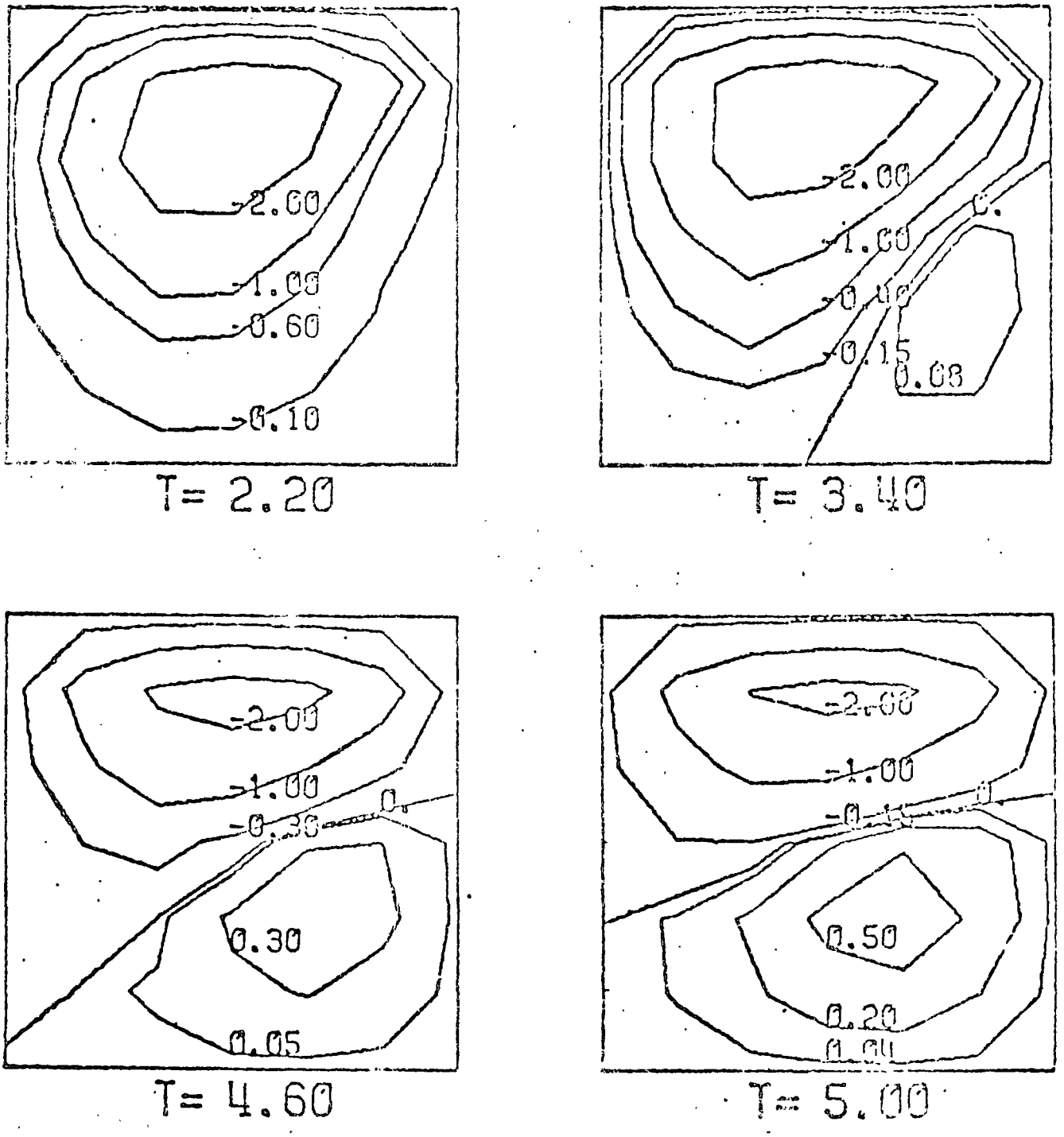
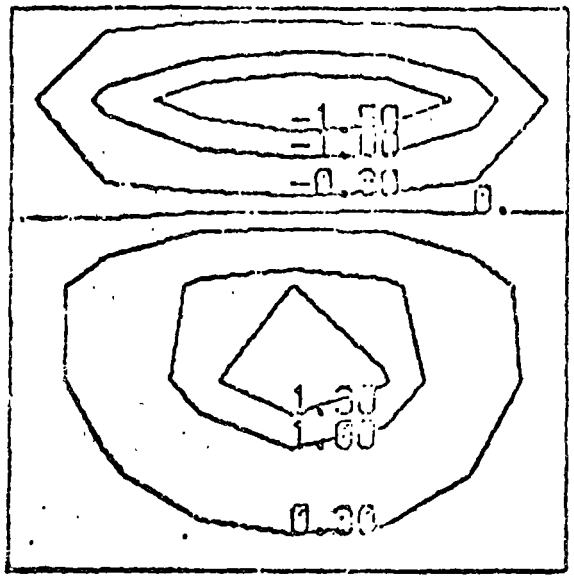
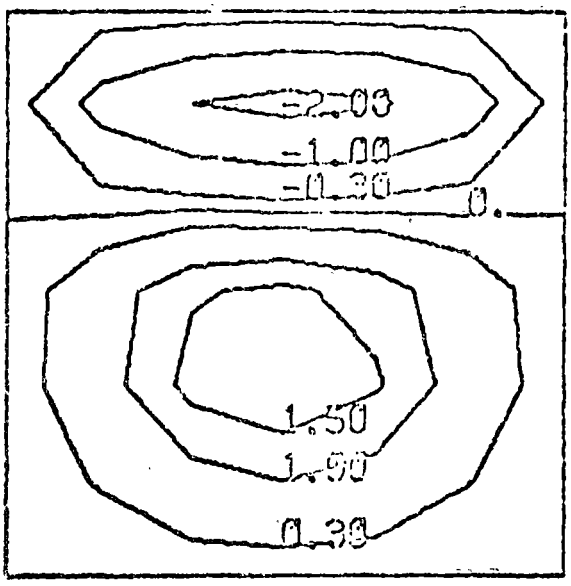


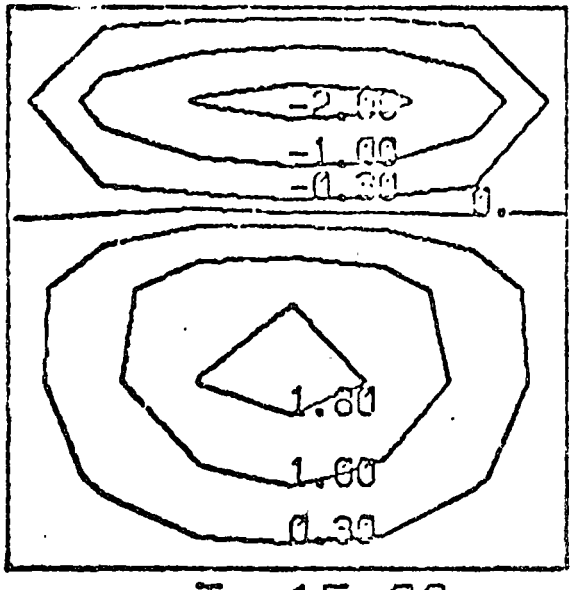
Figure 7. Streamline Pattern for Stratified Flow at Various (Dimensionless) Times



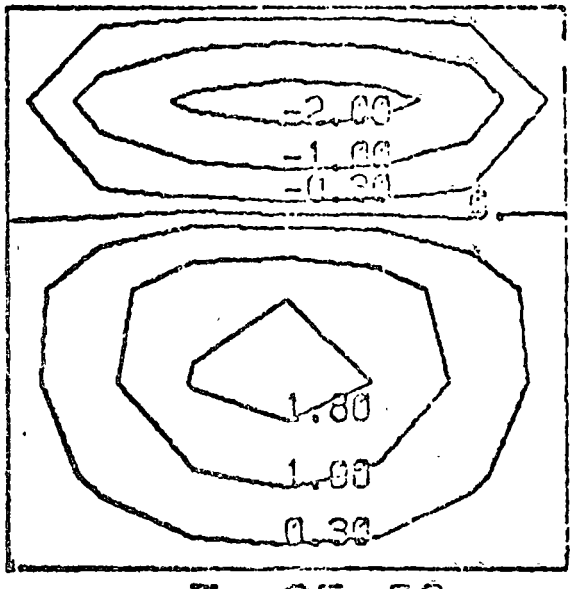
T= 7.00



T= 10.00



T= 15.00



T= 25.00

Figure 7 (continued)

likely to form in the case of a diffuse metalimmion whereas a sharp thermocline promotes two cell circulation. However, the feedback of the currents into the density structure has not been considered, and this feedback may alter the density distribution, thus changing the results of the theory.

Variable viscosity, especially a formulation which is strongly dependent on the density gradient, has a great effect on the ability of the current structure to alter the density distribution. A reduction of viscosity in zones of steep density gradient promotes the oscillations observed previously and increases their amplitude (but does not affect the frequency). The reduction in viscosity also increases the time to equilibrium significantly in those cases in which a steady state exists.

A particularly striking result of the transient calculations is the time scale involved in thermocline formation. A shear applied at the surface may alter the density distribution and create a thermocline-like structure in a few hours versus the weeks involved in the one-dimensional computations. Thus the entire process of the development, maintenance, and erosion of the thermocline is a complex process strongly influenced by the current structure. The "physical constants" (i.e., eddy diffusivity) derived for the one-dimensional analysis have, in reality, little physical meaning when the current structures is neglected.

References

1. Norrie, D. and de Vries, G. The Finite Element Method - Fundamentals and Applications Academic Press, N. Y., 1973.

2. Oden, J. T., Zienkiewicz, O. C., Gallagher, R. H. and Taylor, C. Finite Elements in Flow Problems University of Alabama at Huntsville Press, 1974.

3. Gallagher, R. H., et al Finite Elements in Fluids, V. 1 and 2 J. Wiley, Ltd., 1975.

4. Brebbia, C. and Connor, J. J. Numerical Methods in Fluid Dynamics Pentech Press, London, 1974.

5. Monin, A. S. and Yaglom, A. M. Statistical Fluid Mechanics: Mechanics of Turbulence, Vol. 1, MIT Press, 1971.

6. Liggett, J. A. and Hadjitheodorou, C. "Circulation in Shallow Homogeneous Lakes" Proc. ASCE, J. of the Hyd. Div., 95, No. HY 2, March 1969, pp. 609-620.

7. Gallagher, R. H., Liggett, J. A. and Chan, S.T.K. "Finite Element Circulation Analysis of Variable-Depth Shallow Lakes" Proc. ASCE, J. of the Hyd. Div., 99, No. HY 7, July 1973, pp. 1083-1096.

8. Gallagher, R. H. and Chan, S.T.K. "Higher-Order Finite Element Analysis of Lake Circulation" Computers and Fluids, V. 1, No. 2, 1973, pp. 119-132.

9. Cheng, R. T. "Numerical Investigation of Lake Circulation Around Islands by the Finite Element Method" Int. J. for Numerical Methods in Engrg., 5, No. 1, pp. 103-112, 1972.

10. Tong, Pin "Finite Element Solution of the Wind-Driven Current and its Mass Transport in Lakes" in Numerical Methods in Fluid Dynamics, C. Brebbia and J. J. Connor, Eds., Pentech Press, London, 1974, pp. 440-453.

11. Welander, P. "Wind Action on Shallow Sea, Some Generalizations of Eckmann's Theory" Tellus, 9, No. 1, 1957, pp. 47-52.

12. Loziuk, L., Anderson, J. and Belytschko, T. "Hydrothermal Analysis by the Finite Element Method" Proc. ASCE, J. of Hyd. Div., 98, No. HY 11, November 1972.

13. Loziuk, L., Anderson, J. and Belytschko, T. "Finite Element Approach to Hydrothermal Analysis of Small Lakes" Mtg. Preprint 1799, ASCE National Environmental Meeting, Houston, Texas, October 1972.

14. Liggett, J. A. and Bedford, K. W. "Stratified, Shear-Driven Cavity (Lake) Flow" Proc. ASCE, J. of the Hyd. Div. (To be published).

15. Bedford, K. W. A Numerical Investigation of Stably Stratified Wind Driven Cavity Flow by the Finite Element Method Ph.D. Diss., Cornell University, June 1974.

16. Young, D. L, Liggett, J. A. and Gallagher, R. H. "Steady Stratified Circulation in a Cavity" Proc. ASCE, J. of the Hyd. Div. (To be published).

17. Skiba, E., Unny, T. E. and Weaver, D. S. "A Finite Element Solution for a Class of Two-Dimensional Viscous Fluid Dynamics Problems" Computer Aided Engineering, Proc. Symposium, University of Waterloo, 1971.

18. Debongnie, J. F. "Etude par la Methode des Elements Finis D'un Ecoulement de Stokes Dans un Canal a Paroi Crenelee" Collection des Publications, Faculte des Sciences Appliquees, Universite de Liege, Belgium, No. 48, 1974.

19. Kawahara, M., Yoshimura, N., and Nakagawa, K. "Steady and Unsteady Finite Element Analysis of Incompressible Viscous Flow" Int. J. Num. Meth. in Engrg. (To be published).

20. Zienkiewicz, O. C., Gallagher, R. H. and Hood, P. "Newtonian and Non-Newtonian Viscous Incompressible Flow. Temperature Induced Flows. Finite Element Solutions" Proc. of Conference on the Mathematics of Finite Elements and Applications, Brunel University (To be published).

21. Sundaram, T. R. and Rehm, R. G. "Formation of Maintenance of Thermocline in Temperate Lakes" AIAA Journal, Vol. 9, No. 7, 1970, pp. 1322-1329.

22. Sundaram, T. R., Rehm, R. G., Rudinger, G. and Merritt, G. E. A Study of Some Problems on the Physical Aspects of Thermal Pollution Rept. VT-2790-A-1, Cornell Aero. Lab., Buffalo, N. Y., 1970.

23. Welander, P. "The Thermocline Problem" Phil. Trans Roy. Soc. London Ser. A 270, 1971, pp. 415-421.

24. Connor, J. J. and Wang, J. "Finite Element Modeling of Hydrodynamic Circulation" in Numerical Methods in Fluid Dynamics, Brebbia, C. A. and Connor, J. J., Eds., Pentech Press, London, 1974, pp. 355-387.

25. Taylor, C. and Davis, J. M. "A Finite Element Model of Tides in Estuaries" Finite Element Methods in Flow Problems J. T. Oden, et al, Eds., UAH Press, 1974.

26. Taylor, C. and Davis, J. M. "A Numerical Model of Dispersion in Estuaries" Finite Element Methods in Flows Problems, J. T. Oden, et al, Eds., UAH Press, 1974.

27. Adey, R. A. and Brebbia, C. A. "Finite Element Solution for Effluent Dispersion" in Numerical Methods in Fluid Dynamics, Brebbia, C. A. and Connor, J. J., Eds., Pentech Press, 1974.

28. Pinder, G. F. "Simulation of Ground Water Contamination Using a Galerkin Finite Element Technique" in Numerical Methods in Fluid Dynamics, Brebbia, C. A. and Connor, J. J., Eds., Pentech Press, 1974.

29. Mercer, J. W. and Pinder, G. F. "Finite Element Analysis of Hydrothermal Systems" Finite Element Methods in Flow Problems, J. T. Oden, et al, Eds., UAH Press, 1974.

30. Al Mashidani, G. and Taylor, C. "Finite Element Solutions of the Shallow-Water Equations" in Finite Element Methods in Flow Problems, J. T. Oden, et al, Eds., UAH Press, 1974, pp. 385-398.

31. Young, D. L. Dynamics of Transient Wind-Driven Circulation in a Stably-Stratified Basin by Finite Element Analysis Ph.D. Thesis, Cornell University, Ithaca, N.Y., January 1976.

32. Turner, J. S. Buoyancy Effects in Fluids Cambridge Univ. Press, London, 1973.

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION
Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS

STABILITY OF FINITE ELEMENT MODELS FOR CONVECTION-DIFFUSION

DR. JEROME J. CONNOR

MAYO, 1978.

STABILITY OF FINITE ELEMENT MODELS FOR CONVECTION-DIFFUSION

by

Jerome J. Connor[1] and George C. Christodoulou[2]

INTRODUCTION

The convection-diffusion equation, a typical parabolic partial differential equation, is of great importance in a variety of fields. It is essentially an expression of conservation of a quantity which is subject to advection and at the same time spreads out due to molecular or larger scale mechanisms. Applications are in heat conduction, flow in porous media, vorticity transport in viscous flows, and dispersion in air or water bodies.

If the domain is irregular, the equation has to be solved numerically and criteria for the stability of the particular numerical scheme used is essential. A large number of finite difference techniques have been employed in the past and their stability and accuracy characteristics have been examined for simple convection, simple diffusion, and combined convection-diffusion. A review of these methods can be found in [7].

Use of the finite element method in fluid problems has recently become quite popular [2,3]. The finite element discretization is normally applied only in the spatial domain, while time integration proceeds through conventional schemes. Rigorous

[1]Professor, Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

[2]Assistant Professor, Applied Hydraulics Lab., National Technical University of Athens, Athens, Greece

theoretical analysis has focused on such fundamental problems as the convergence of the finite element approximation to the true solution, determination of error bounds, etc. [10], but has failed so far to yield practical results. This is due largely to the inherent difficulty in extracting simple expressions in terms of the problem parameters from the matrix equations defining the stability limit. Thus, important issues in problem-solving, such as the selection of time step, are mostly dealt with through experience or experimentation.

This paper discusses a stability investigation of the finite element method applied to the 2-D convection-diffusion equation. A generalized (arbitrary) spatial discretization is assumed, along with a simple implicit iterative scheme based on the trapezoidal rule for time integration. The method is shown to be unconditionally stable for an arbitrary grid, constant parameters, and no iteration, except under certain rare boundary conditions. General criteria for convergence of the iteration procedure are developed and specialized for the particular case of triangular elements with linear interpolation functions. The effect of the finite element discretization on the accuracy of the solution is also briefly examined. Finally, results of numerical experiments which confirm the theoretical results are presented.

FINITE ELEMENT FORMULATION

The general convection-diffusion equation has the form

$$\frac{\partial C}{\partial t} + \vec{\nabla} \cdot (\vec{U} C) = \vec{\nabla} \cdot \vec{Q} - f_d - f_s \qquad (1)$$

where C is the solution variable, e.g. the concentration of a constituent, $\vec{U}$ is the (local) velocity vector, $\vec{Q}$ is the diffusive flux vector, f_d and f_s are the decay and source terms, respectively. In the simplest case of a linear decay, f_d is expressed as

$$f_d = k\,C \tag{2}$$

where k is the decay constant (> 0).

A two-dimensional domain is considered here. An analogous procedure can be applied to one dimensional problems. Incorporating (2) into (1), considering the fluid incompressible, and integrating through the transverse direction leads to an "averaged" equation,

$$\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} + v\frac{\partial C}{\partial y} = -\frac{\partial}{\partial x}Q_x + \frac{\partial}{\partial y}Q_y - kC + f_s \tag{3}$$

where u, v are the average velocity components and Q_x, Q_y are the average diffusive fluxes in the x, y direction. If the diffusion mechanism is assumed to follow a Fickian behavior, one can write

$$\begin{Bmatrix} Q_x \\ Q_y \end{Bmatrix} = -\begin{bmatrix} E_{xx} & E_{xy} \\ E_{yx} & E_{yy} \end{bmatrix} \begin{Bmatrix} \partial C/\partial x \\ \partial C/\partial y \end{Bmatrix} \tag{4}$$

where the set of diffusion coefficients comprise a second order symmetrical tensor $\underset{\sim}{E}$. Finally the boundary conditions for the problem are of two types (see Figure 1):

(i) Essential, i.e., concentration specified

$$C = C^* \quad \text{on} \quad S_c \tag{5a}$$

or

(ii) Natural, i.e., normal diffusive flux specified

$$Q_n = Q_n^* \quad \text{on} \quad S_q \tag{5b}$$

Applying the method of weighted residuals to (3) and integrating the flux terms by parts leads to the symmetrical weak form [2]:

$$R = \iint_A \left(\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} + v\frac{\partial C}{\partial y} + kC - f_s \right) W dA$$

$$+ \iint \left(-Q_x \frac{\partial W}{\partial x} + Q_y \frac{\partial W}{\partial y} \right) dA + \int_{S_q} Q_n^* W dS = 0 \tag{6}$$

where W represents a weighting function. Since only first derivatives appear in this form (when Q_x, Q_y are expressed through Eq. (4)), the trial function C (approximate solution) and weighting function W have to be only piecewise continuous within the domain A. At the boundary, the trial function is required to satisfy (5a), while the weighting function must satisfy the homogeneous form, i.e., W = 0 on S_c.

In the finite element method, the domain is subdivided into "elements" and the total residual, R, which is required by Eq. (6) to vanish, is evaluated as the sum of the element residuals, R^e. The trial and test variables distributions over an element domain are expressed as

$$C = \underset{\sim}{N}\, \underset{\sim}{C}^e$$

$$W = \underset{\sim}{N}\, \underset{\sim}{W}^e \tag{7}$$

where $\underset{\sim}{C}^e$, $\underset{\sim}{W}^e$ are column vectors containing the values at the element nodes, and $\underset{\sim}{N}$ contains interpolation functions restricted only by interelement continuity of C and W. Using Eq. (7), and (4), the residual for a single element expands to

$$R^e = (\underset{\sim}{W}^e)^T \left\{ \left(\iint \underset{\sim}{N}^T \underset{\sim}{N}\, dA \right) \frac{\partial \underset{\sim}{C}^e}{\partial t} + \left(\iint \underset{\sim}{N}^T \left(u \frac{\partial \underset{\sim}{N}}{\partial x} + v \frac{\partial \underset{\sim}{N}}{\partial y} \right) dA \right) \underset{\sim}{C}^e \right.$$
$$+ k \left(\iint \underset{\sim}{N}^T \underset{\sim}{N}\, dA \right) \underset{\sim}{C}^e + \left(\iint \left[\frac{\partial \underset{\sim}{N}^T}{\partial x} \quad \frac{\partial \underset{\sim}{N}^T}{\partial y} \right] \underset{\sim}{E} \begin{Bmatrix} \frac{\partial \underset{\sim}{N}}{\partial x} \\ \frac{\partial \underset{\sim}{N}}{\partial y} \end{Bmatrix} dA \right) \underset{\sim}{C}^e$$
$$\left. - \iint \underset{\sim}{N}^T f_s\, dA + \int_{S_\phi} \underset{\sim}{N}^T Q_n^*\, dS \right\} \tag{8}$$

Introducing simplified notation for the matrix coefficients results in

$$R^e = (\underset{\sim}{W}^e)^T \left\{ \underset{\sim}{M}^e \dot{\underset{\sim}{C}}^e + (\underset{\sim}{A}^e + \underset{\sim}{K}^e + \underset{\sim}{D}^e) \underset{\sim}{C}^e + \int \underset{\sim}{N}^T Q_n^*\, dS - \iint \underset{\sim}{N}^T f_s\, dA \right\} \tag{9}$$

where the various element matrices are defined as

$$\underset{\sim}{M}^e = \iint \underset{\sim}{N}^T \underset{\sim}{N}\, dA \qquad \text{(geometrical)}$$

$$\underset{\sim}{A}^e = \iint \underset{\sim}{N}^T \left(u \frac{\partial \underset{\sim}{N}}{\partial x} + v \frac{\partial \underset{\sim}{N}}{\partial y} \right) dA \qquad \text{(advection)}$$

$$\underset{\sim}{K}^e = \iint \partial \underset{\sim}{N}^T \underset{\sim}{E}\, \partial \underset{\sim}{N}\, dA \qquad \text{(diffusion)}$$

$$\underset{\sim}{D}^e = k \iint \underset{\sim}{N}^T \underset{\sim}{N}\, dA \qquad \text{(decay)} \tag{10}$$

Both $\underset{\sim}{M}^e$ and $\underset{\sim}{D}^e$ are symmetric positive definite matrices. The character of $\underset{\sim}{K}^e$ depends on $\underset{\sim}{E}$. Since diffusion is normally a "dissipative" mechanism, it is reasonable to assume $\underset{\sim}{E}$ is positive definite. Then, $\underset{\sim}{K}^e$ is positive semi-definite with respect to $\underset{\sim}{C}^e$.

Summing up the element residuals, one obtains the total set of equations relating the nodal concentration.

$$\underset{\sim}{M}\,\dot{\underset{\sim}{C}} + (\underset{\sim}{A} + \underset{\sim}{K} + \underset{\sim}{D})\,\underset{\sim}{C} = \text{contributions from sources and boundary conditions} \tag{11}$$

The symmetry and definiteness properties of $\underset{\sim}{M}$, $\underset{\sim}{D}$ and $\underset{\sim}{K}$ are the same as for the corresponding element matrices.

To establish the characteristics of the advection matrix, partial integration is carried out over the element domain

$$\underset{\sim}{A}^e\underset{\sim}{C}^e = \left(\iint \underset{\sim}{N}^T\left(u\frac{\partial \underset{\sim}{N}}{\partial x} + v\frac{\partial \underset{\sim}{N}}{\partial y}\right)dA\right)\underset{\sim}{C}^e$$

$$= \left(\oint_{S^e} \underset{\sim}{N}^T u_n \underset{\sim}{N}\, dS\right)\underset{\sim}{C}^e - \left(\iint\left(\frac{\partial \underset{\sim}{N}^T}{\partial x}u\underset{\sim}{N} + \frac{\partial \underset{\sim}{N}^T}{\partial y}v\underset{\sim}{N}\right)dA\right)\underset{\sim}{C}^e$$

$$= \left(\oint_{S^e} \underset{\sim}{N}^T u_n \underset{\sim}{N}\, dS\right)\underset{\sim}{C}^e - \left(\underset{\sim}{A}^e\right)^T\underset{\sim}{C}^e \tag{12}$$

where u_n denotes the outward normal velocity. When the element residuals are summed, the interior element line integrals will cancel out and provided that continuous expansions are used for u, v and C. Then,

$$\underset{\sim}{A}\,\underset{\sim}{C} = \left(\int_{S_u} \underset{\sim}{N}^T u_n \underset{\sim}{N}\, dS\right)\underset{\sim}{C} - \underset{\sim}{A}^T\underset{\sim}{C} \tag{13}$$

The line integral in Eq. (13) is restricted only to that part of the boundary, S_u, where the normal velocity is finite. This excludes land boundaries which, by definition, require $u_n=0$. If $S_u=0$ or the concentration is set to zero on S_u, the integral vanishes and the system advection matrix is skew-symmetric.

In general, one can write:

$$\underset{\sim}{A} = \underset{\sim}{A}_s + \underset{\sim}{A}_{ss} \tag{14}$$

where $\underset{\sim}{A}_s$, $\underset{\sim}{A}_{ss}$ denote the symmetric and skewsymmetric parts of $\underset{\sim}{A}$, respectively. According to (13) $\underset{\sim}{A}_s$ is associated with the boundary segments on which the normal velocity and concentration are finite.

No use has been made of the particular form of the interpolation functions N, or the element shape. Therefore, the conclusions as to the properties of the system matrices are valid for an arbitrary spatial discretization based on continuous expansions ,

STABILITY OF TIME INTEGRATION

One of the simplest time integration methods is based on the trapezoisal rule. In order to examine its stability characteristics when applied to the discretized finite element transport Equation, (11), the homogeneous form of the latter will be considered, i.e.,

$$\underset{\sim}{M}\,\dot{\underset{\sim}{C}} + \left(\underset{\sim}{A} + \underset{\sim}{K} + \underset{\sim}{D}\right)\underset{\sim}{C} = \underset{\sim}{O} \tag{15}$$

The recurrence relation for the trapezium method is

$$\left[\underset{\sim}{M} + \frac{\Delta t}{2}\left(\underset{\sim}{A}+\underset{\sim}{K}+\underset{\sim}{D}\right)_{m+1}\right]\underset{\sim}{C}_{m+1} = \left[\underset{\sim}{M} - \frac{\Delta t}{2}\left(\underset{\sim}{A}+\underset{\sim}{K}+\underset{\sim}{D}\right)_{m}\right]\underset{\sim}{C}_{m} \tag{16}$$

where the subscript denotes the time index.

Assuming the advection, dispersion, and decay matrices are time independent, one can express the solution of (16) as

$$\underset{\sim}{C}_{n+1} = \underset{\sim}{B}\, C_n \tag{17}$$

where $\underset{\sim}{B}$ is the amplification matrix. An alternate form is

$$\underset{\sim}{C}_n = \lambda^n \underset{\sim}{\phi} \tag{18}$$

where λ, $\underset{\sim}{\phi}$ are the eigenvalue, eigenvector of $\underset{\sim}{B}$. The sufficient condition for stability is

$$\|\underset{\sim}{B}\| < 1 \quad \Rightarrow \quad |\lambda_{max}| < 1 \tag{19}$$

Substituting (18) into (16), premultiplying both sides by $\bar{\underset{\sim}{\phi}}^T$, the transpose of the complex conjugate of $\underset{\sim}{\phi}$, one obtains:

$$\lambda\, \bar{\underset{\sim}{\phi}}^T \left[\underset{\sim}{M} + \frac{\Delta t}{2}\left(\underset{\sim}{A} + \underset{\sim}{K} + \underset{\sim}{D}\right)\right]\underset{\sim}{\phi} = \bar{\underset{\sim}{\phi}}^T \left[\underset{\sim}{M} - \frac{\Delta t}{2}\left(\underset{\sim}{A} + \underset{\sim}{K} + \underset{\sim}{D}\right)\right]\underset{\sim}{\phi} \tag{20}$$

Since $\underset{\sim}{M}$, $\underset{\sim}{K}$ and $\underset{\sim}{D}$ are symmetric positive matrices, it follows that

$$\begin{aligned} \bar{\underset{\sim}{\phi}}^T \underset{\sim}{M}\, \underset{\sim}{\phi} &= m > 0 \\ \bar{\underset{\sim}{\phi}}^T \underset{\sim}{K}\, \underset{\sim}{\phi} &= \varepsilon > 0 \\ \bar{\underset{\sim}{\phi}}^T \underset{\sim}{D}\, \underset{\sim}{\phi} &= d > 0 \end{aligned} \tag{21}$$

The advection term expands to

$$\bar{\underset{\sim}{\phi}}^T \underset{\sim}{A}\, \underset{\sim}{\phi} = a_s + i\, a_{ss} \tag{22}$$

where a_s is related to the symmetric and a_{ss} to the skew-symmetric part of $\underset{\sim}{A}$. With this notation, one can write

$$\lambda = \frac{m - \frac{\Delta t}{2}\left(a_s + \varepsilon + d\right) - i\,\frac{\Delta t}{2}\,a_{ss}}{m + \frac{\Delta t}{2}\left(a_s + \varepsilon + d\right) + i\,\frac{\Delta t}{2}\,a_{ss}} \tag{23}$$

If $a_s = 0$, $|\lambda| < 1$ for abritrary Δt.
That is, the integration scheme is unconditionally stable for an arbitrary grid when the system parameters are constant.

The value of a_s can be different than zero only if there is a segment of S_q on which the normal velocity does not vanish. Noting (12), and the positive definite property of $\underset{\sim}{N}^T\underset{\sim}{N}$,

$$a_s = \underset{\sim}{\bar{\phi}}^T \underset{\sim}{A}_s \underset{\sim}{\phi} = \underset{\sim}{\phi}^T \left(\int_{S_q} u_n \underset{\sim}{N}^T \underset{\sim}{N}\, dS \right) \underset{\sim}{\phi} \tag{24}$$

one finds that a_s has the sign of u_n. Therefore, when the normal velocity is directed outwards, $a_s > 0$, and Eq. (23) indicates that $|\lambda| < 1$. Actually, the stability is enhanced in this case. However, when u_n is directed inward, $a_s < 0$, and the stability depends on the relative magnitudes of a_s and $\varepsilon + d$. Sufficiently strong diffusion or decay mechanisms can offset a negative a_s.

It is a common practice, based on physical considerations, to specify the concentration during inflow and the concentration gradient during outflow. With this procedure, stability of the

scheme is usually maintained. The source of instability for the case of inflow from a boundary with prescribed concentration gradient is the introduction into the domain of an uncontrolled quantity of material.

ITERATION CONVERGENCE

Although the feature of unconditional stability makes the trapezoidal integration scheme extremely attractive, problems involving time variability of parameters or inputs will necessarily require a restriction on the time step. In such a case it would be economical to invert a new matrix of the form $\underset{\sim}{M} + \frac{\Delta t}{2}(\underset{\sim}{A}+\underset{\sim}{K}+\underset{\sim}{D})$ at every time step only for a very small problem. An iteration procedure is, therefore, preferred. Iteration has the additional advantage of being able to handle nonlinear decay, nonfickian diffusion, etc. For the homogeneous problem examined here, the recurrence relation is expressed as

$$\underset{\sim}{M}\,\underset{\sim}{C}_{m+1}^{(i+1)} = \left[\underset{\sim}{M} - \frac{\Delta t}{2}\left(\underset{\sim}{A}+\underset{\sim}{K}+\underset{\sim}{D}\right)_{m}\right]\underset{\sim}{C}_{m} - \frac{\Delta t}{2}\left(\underset{\sim}{A}+\underset{\sim}{K}+\underset{\sim}{D}\right)_{m+1}\underset{\sim}{C}_{m+1}^{(i)} \tag{25}$$

where the superscript denotes the iteration index.

The sufficient condition for convergence is the requirement that the norm of the amplification matrix be less than unity.

$$\left\| \underset{\sim}{M}^{-1} \frac{\Delta t}{2}\left(\underset{\sim}{A}+\underset{\sim}{K}+\underset{\sim}{D}\right)_{m+1} \right\| < 1 \tag{26}$$

Solving for Δt yields

$$\Delta t < \frac{2}{\left\| \underset{\sim}{M}^{-1}\left(\underset{\sim}{A}+\underset{\sim}{K}+\underset{\sim}{D}\right)\right\|} \tag{27}$$

A more conservative form is

$$\Delta t < \frac{2}{\|\underset{\sim}{M}^{-1}\underset{\sim}{A}\| + \|\underset{\sim}{M}^{-1}\underset{\sim}{K}\| + \|\underset{\sim}{M}^{-1}\underset{\sim}{D}\|} \tag{28}$$

In principle, the norm expression (27), can be evaluated. However, calculation of the commonly used eigenvalue norms involves long machine computations once the matrices are formed. An explicit relation between the time step and the parameters of the problem would be very desirable for practical applications. An approximate relation can be derived by evaluating the norm expressions for an individual element. Provided there are no drastic changes in the grid or the prameters over the domain, the conclusions reached at the element level can be generalized for the whole system. That is, satisfaction of (27) for the "worst" element would indicate convergence for the total system of elements.

CRITERIA FOR LINEAR TRIANGLES

Norms for the triangular element (see Figure 2) and a linear expansion are presented in this section. Starting with the expansion

$$\underset{\sim}{N} = \begin{bmatrix} \xi_1 & \xi_2 & \xi_3 \end{bmatrix}$$

$$\frac{\partial \underset{\sim}{N}}{\partial x} = \frac{1}{2A^e}\begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} = \frac{1}{2A^e}\,\underset{\sim}{b}$$

$$\frac{\partial \underset{\sim}{N}}{\partial y} = \frac{1}{2A^e}\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} = \frac{1}{2A^e}\,\underset{\sim}{a} \tag{29}$$

where A^e is the area of the triangle and

$$a_1 = x_3 - x_2 \qquad b_1 = y_2 - y_3$$

$$a_2 = x_1 - x_3 \qquad b_2 = y_3 - y_1$$

$$a_3 = x_2 - x_1 \qquad b_3 = y_1 - y_2 \tag{30}$$

one obtains the following element matrices:

(i) Geometrical matrix

$$\underset{\sim}{M} = \frac{A}{12}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

$$\underset{\sim}{M}^{-1} = \frac{3}{A}\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix} = \frac{12}{A}\left(\underset{\sim}{I} - \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\right) \tag{31}$$

(ii) Advection matrix, for uniform flow (u,v),

$$\underset{\sim}{A} = \frac{1}{24}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}\left(\underset{\sim}{u}^e \underset{\sim}{b} + \underset{\sim}{v}^e \underset{\sim}{a}\right)$$

$$= \frac{U}{2A}\,\underset{\sim}{M}\begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix}\begin{bmatrix} \Delta s_1 \cos\phi_1 & \Delta s_2 \cos\phi_2 & \Delta s_3 \cos\phi_3 \end{bmatrix} \tag{32}$$

where $U = \sqrt{u^2+v^2}$ is the magnitude of the velocity vector, Δs_i is the length of the ith side, and ϕ_i is the angle between the velocity vector and the inward normal to Δs_i. This notation is illustrated in Figure 3.

(iii) Diffusion matrix, for isotropic conditions

($E_{xx} = E_{yy} = E$, $E_{xy} = E_{yx} = 0$).

$$\underset{\sim}{K} = \frac{E}{4A}\begin{bmatrix} \Delta s_1^2 & -\Delta s_1 \Delta s_2 \cos\theta_3 & -\Delta s_1 \Delta s_3 \cos\theta_2 \\ & \Delta s_2^2 & -\Delta s_2 \Delta s_3 \cos\theta_1 \\ \text{Sym} & & \Delta s_3^2 \end{bmatrix}$$

where θ_i is the angle opposite side i.

(iv) Decay matrix

$$\underset{\sim}{D} = k \iint \underset{\sim}{N}^T \underset{\sim}{N} \, dA = k \underset{\sim}{M} \tag{34}$$

Proceeding now to form the products appearing in (28), it is first noticed that

$$\underset{\sim}{M}^{-1} \underset{\sim}{D} = k \tag{35}$$

Since the rows of $\underset{\sim}{K}$ are linearly dependent, and noting the second form of $\underset{\sim}{M}^{-1}$, one can write

$$\underset{\sim}{M}^{-1} \underset{\sim}{K} = \frac{12}{A} \underset{\sim}{K} \tag{36}$$

Specializing (36) for the particular case of an equilateral triangle

$$\underset{\sim}{M}^{-1} \underset{\sim}{K} = \frac{12}{A} \frac{E \Delta S^2}{8A} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \tag{37}$$

The matrix eigenvalues are 0, 3, 3. Using 3 yields a lower bound on Δt,

$$\| \underset{\sim}{M}^{-1} \underset{\sim}{K} \| = \frac{12 E \Delta S^2}{8 \left(\frac{3}{16} \Delta S^4 \right)} \cdot 3 = 24 \frac{E}{\Delta S^2} \tag{38}$$

This result is too conservative. A more reasonable estimate is obtained with the "average" value of the eigenvalues, i.e., 2.

$$\| \underset{\sim}{M}^{-1} \underset{\sim}{K} \|_{av} = 16 \frac{E}{\Delta S^2} \tag{39}$$

The last term, which involves $\underset{\sim}{M}^{-1}$ and $\underset{\sim}{A}$ has the form

$$\underset{\sim}{M}^{-1}\underset{\sim}{A} = \frac{U}{2A}\begin{Bmatrix}1\\1\\1\end{Bmatrix}\begin{bmatrix}\Delta S_1 \cos\phi_1 & \Delta S_2 \cos\phi_2 & \Delta S_3 \cos\phi_3\end{bmatrix} \tag{40}$$

A norm measure is generated using the Eudidean norms for the vectors,

$$\|\underset{\sim}{M}^{-1}\underset{\sim}{A}\| < \frac{U}{2A}\sqrt{3}\left\{(\Delta S_1 \cos\phi_1)^2 + (\Delta S_2 \cos\phi_2)^2 + (\Delta S_3 \cos\phi_3)^2\right\}^{1/2} \tag{41}$$

Specializing for an equilateral triangle,

$$\|\underset{\sim}{M}^{-1}\underset{\sim}{A}\| < 2\sqrt{\frac{3}{2}}\cdot\frac{U}{\Delta S} \tag{42}$$

Similar expressions can be derived for any triangular or rectangular shape. The above expressions also apply for a right triangle, provided the flow is parallel to one of the short sides.

Combining (25), (39), and (42), one obtains an estimate for stability of the iteration scheme applied to an equilateral grid,

$$\Delta t < \frac{1}{1.2\frac{U}{\Delta S} + 8\frac{E}{\Delta S^2} + \frac{k}{2}} \tag{43}$$

In actual applications, oen designs the grid with approximately equilateral triangles, avoiding angles in excess of 90°. Equation (43) provides a good starting point for selecting the time step in any given problem. Its validity is tested experimentally in a later section. Its primary value, irrespective of the numerical constants which vary with element type,

lies in the inclusion of all relevant parameters in a single expression. In the past, these have been examined separately, if at all, and only by numerical experimentation. For instance, [5] suggests that for satisfactory time integration using the same iterative scheme and the same type of elements, both of the following conditions must hold:

$$\Delta t < \frac{\Delta s}{10\,U} \qquad \Delta t < \frac{\Delta s^2}{10\,E} \tag{44}$$

These bounds are stricter than (43) especially with respect to advection. In problems with significant spatial variations, it is clear that the largest values of $E/\Delta s^2$ and $U/\Delta s$ limit the time step. As a consequence, a local refinement of the grid for better resolution will lead to a smaller allowable time step.

ACCURACY

In addition to stability considerations, an important issue that has to be addressed is the accuracy of the numerical solution i.e., how close the true solution is being approximated. Accuracy depends on the space and time discretization and the type of problem being solved. The two basic errors considered in diffusion problems are those of numerical damping and numerical dispersion. The former relates to excessive (or inadequate) damping of the magnitude of individual wave components and is sometimes expressed by an artificial diffusion coefficient.

The latter relates to incorrect phase propagation of individual wave components, which leads to distortion of the shape of the overall distribution. These errors may be easily examined in finite difference schemes; they can also be examined, in an analogous way, when the finite element discretization involves a regular grid. Thus, it has been found that the trapezoidal integration scheme with linear triangular elements gives negligible phase error and rather small amplitude error [1] - actually being neutrally stable for simple convection. Similar results have been obtained for rectangular isoparametric elements [9].

However, when significant high frequency components exist in the function to be approximated, serious difficulties may arise from the inability of the grid to represent them adequately. In particular, linear interpolation functions are suitable for describing the Gaussian or exponential analytical solutions of diffusion problems only if the grid is fine. Higher order interpolation is superior, but also more expensive. The usual result of the inadequacy of the grid to accommodate steep gradients is the appearance of spatial oscillations and negative concentrations in the numerical solution. This unnatural behavior is due basically to the "coarseness" of the spatial discretization and it has been shown that it would occur even if the problem is one of steady state [7].

Considering a regular finite element grid (Figure 4) and a one-dimensional steady state problem, the discretized equation for Node A is [1]:

$$\frac{U\Delta S}{6}\left(C_B + 2C_C + C_D - C_E - 2C_F - C_G\right) + E\left(2C_A - C_C - C_E\right) = 0 \qquad (45)$$

Suppose that, due to the presence of a source and because the problem is one-dimensional,

(i) the concentrations at C and D are relatively high,

$$C_C \simeq C_D = M.$$

(ii) nodes F and G are essentially out of the plume,

$$C_F \simeq C_G \simeq 0.$$

(iii) $C_B \simeq C_E \simeq C_A$.

Specializing (45) yields

$$C_A \simeq \frac{M}{2}\left(1 - \frac{U\Delta S}{2E}\right) \tag{46}$$

This shows that, in order to avoid negative concentrations upstream of a continuous source, the following condition has to be satisfied:

$$\frac{E}{U\Delta S} > \frac{1}{2} \tag{47}$$

Condition (47) is analogous to the restriction on grid Reynolds number required in central finite difference schemes [7]. Its applicability to the finite element discretization has been established earlier, through numerical experiments [5]. Violation of (47) is sufficient to cause negative concentrations for continuous source 1-D problems. Applying a similar type of approximate analysis as above has indicated that (47) is somewhat conservative for 1-D transient problems, but is not quite adequate for 2-D problems [1]. The additional difficulty that arises in a two-dimensional domain is the singularity of the

analytical solution for a continuous point source. Local refinement of the grid and spreading of the source over a few elements are the easiest ways to improve the numerical solution. If the immediate vicinity of the source is to be modeled accurately, inclusion of singular terms in the trial functions should be considered [8].

EXPERIMENTAL RESULTS

Condition (43) limiting the time step,

$$1.2\,\frac{U\Delta t}{\Delta s} + 8\,\frac{E\Delta t}{\Delta s^2} + 0.5\,k\Delta t < 1$$

is written in the more general form

$$\mu_1\,\frac{U\Delta t}{\Delta s} + \mu_2\,\frac{E\Delta t}{\Delta s^2} + \mu_3\,k\Delta t < 1 \qquad (48)$$

where μ_1, μ_2, μ_3 are numerical constants dependent on the type of the elements used. If the nondimensional groups, $U\Delta t/\Delta s$, $E\Delta t/\Delta s^2$, $k\Delta t$ are viewed as Cartesian coordinates, a useful geometrical interpretation of (48) emerges: the sufficient condition corresponds to bounding the "acceptable" space by a certain plane surface. In the absence of decay, which typically gives a negligible contribution, the space is reduced to two-dimensions. The inequality

$$1.2\,\frac{U\Delta t}{\Delta s} + 8\,\frac{E\Delta t}{\Delta s^2} < 1 \qquad (49)$$

determines a theoretically "safe" area bound by a straight line, as shown in Figure 4. A less conservative inequality [1],

$$\left(1.2\ \frac{U\Delta t}{\Delta s}\right)^2 + \left(\frac{\vartheta E \Delta t}{\Delta s}\right)^2 < 1 \qquad (50)$$

defines an elliptical boundary.

To test the validity of the theoretical results, a large number of runs were carried out using the one-dimensional grid shown in Figure 5. A point source was simulated by loading the three nodes marked with dots. Most runs involved continuous releases, but instantaneous injections were also made. The contribution of the decay term was generally less than 5%, and thus neglected. For each run the corresponding point was plotted on Figure 4. The symbols used to clarify the runs with respect to iteration convergence behavior and the occurence of significant spatial oscillations are explained in Table 1.

The most important feature of Figure 4 is that all runs which exhibit difficulty with iteration convergence lie outside the "safe" region. Not too far from the elliptical boundary, there are points representing runs that rapidly become unstable. Points closer to that boundary, but still outside, generally have iteration errors of 20 to 75%, sometimes decreasing slowly over time. Since there is a limit of 10 iterations per time step in the program, it is not known whether these runs would eventually become unstable if allowed to continue iterating. Apparently, when the iteration is stopped with s small error, the behavior tends to improve over the following time steps. Of course, these errors are accumulated in the solution.

Runs between the two boundaries defined by (49) and (50) have errors of less than 10% which diminish with time. Finally, runs within the inner boundary generally converge very well, with errors less than 1%.

It may be concluded that the theoretical criteria, indended as sufficient conditions, are indeed quite appropriate as such. The elliptical bound (50) is not too conservative, in view of the relatively large errors occurring outside its domain. The linear boundary (49) is somewhat conservative far from the axes; this is a consequence of working with (28) rather than the actual condition (27). Numerical experiments for 2 D problems and applications involving irregular grids of natural water bodies [1] have further confirmed the validity of the theoretical criteria.

The other important result of the experiments is associated with the accuracy condition (47). It is seen that the line $E/U\Delta s = 1/2$ differentiates the regions where runs do or do not show appreciable upstream negative concentrations and spatial oscillations. These oscillations become more severe near the $U\Delta t/\Delta s$ axis, as the ratio $E/U\Delta s$ diminishes, and they are practically eliminated as $E/U\Delta s$ increases slightly above 1/2.

Accuracy considerations significantly reduce the area of acceptability of combinations ($U\Delta t/\Delta s$, $E\Delta t/\Delta s^2$) to a much smaller set than that required for iteration convergence. Fortunately, continuous source problems, which are the most

demanding, are not too sensitive to the value of the diffusion coefficient; this may be changed by almost an order of magnitude without appreciable change in the results [4]. Therefore, it may be possible in some cases to improve the numerical solution by artificially increasing the value of the diffusion coefficient. Another alternative is to resort to higher order elements.

CONCLUSIONS

It has been shown that the trapezoidal integration scheme applied to the discretized convection-diffusion equation (including decay) is unconditionally stable for an arbitrary grid and constant system parameters. This was based upon the examination of the character of the matrices involved, in particular, the skewsymmetry of the advection matrix.

When the system parameters are time dependent, an iterative solution technique is preferred. Its convergence requirements imply some restriction on the time step. Conservative bounds on the time step have been developed for the case of linear triangular elements, based upon a simplified analysis at the individual element level. Results of numerical experiments, mostly on a 1-D grid, confirm to a large extent the theoretical arguments.

Accuracy considerations, related to oscillations of the solution, limit rather severely the use of linear elements in some practical applications. Resort to higher order elements may be worthwhile, despite the increased cost, when there is weak diffusion.

Figure Captions

Figure 1 Solution Field

Figure 2 Typical Linear Triangular Element

Figure 3 Regular Finite Element Grid

Figure 4 Comparison of 1-D Trial Runs with Theoretical Criteria

Figure 5 1-D Test Grid

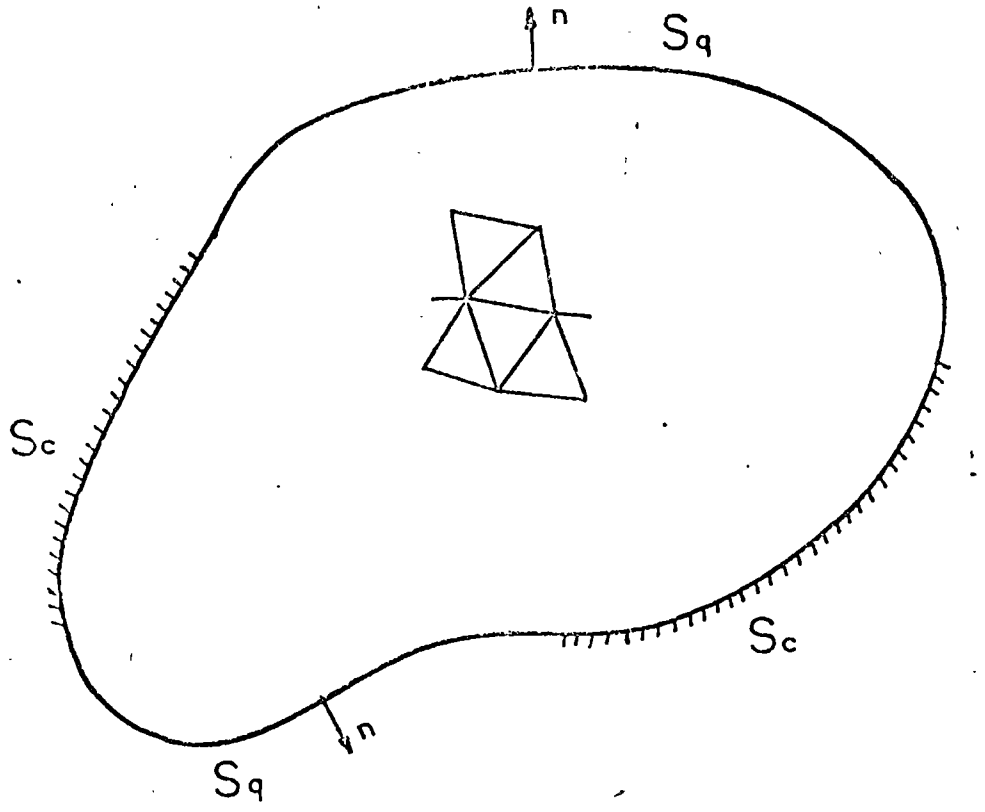


Figure 1 Solution Field

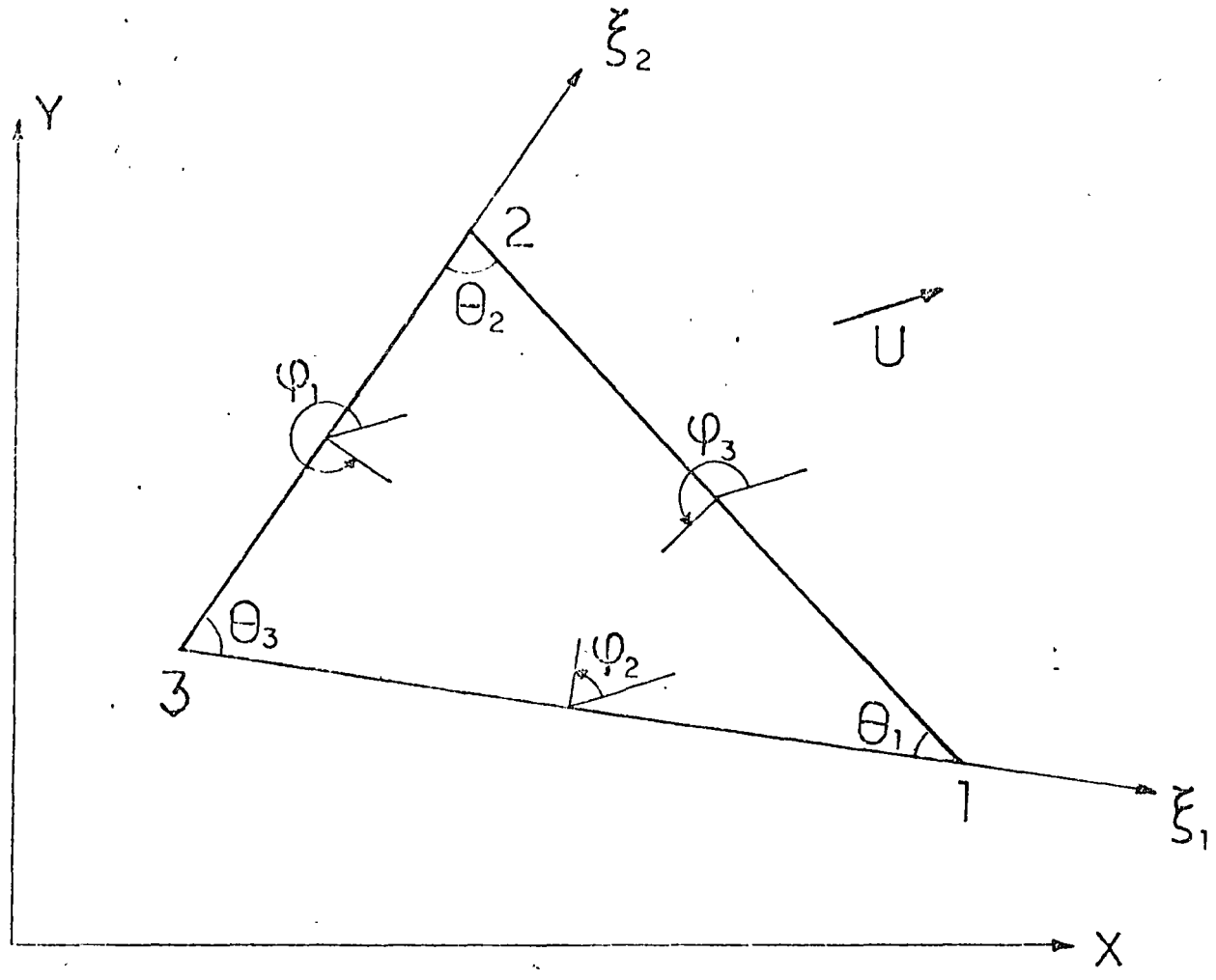


Figure 2 Typical Linear Triangular Element

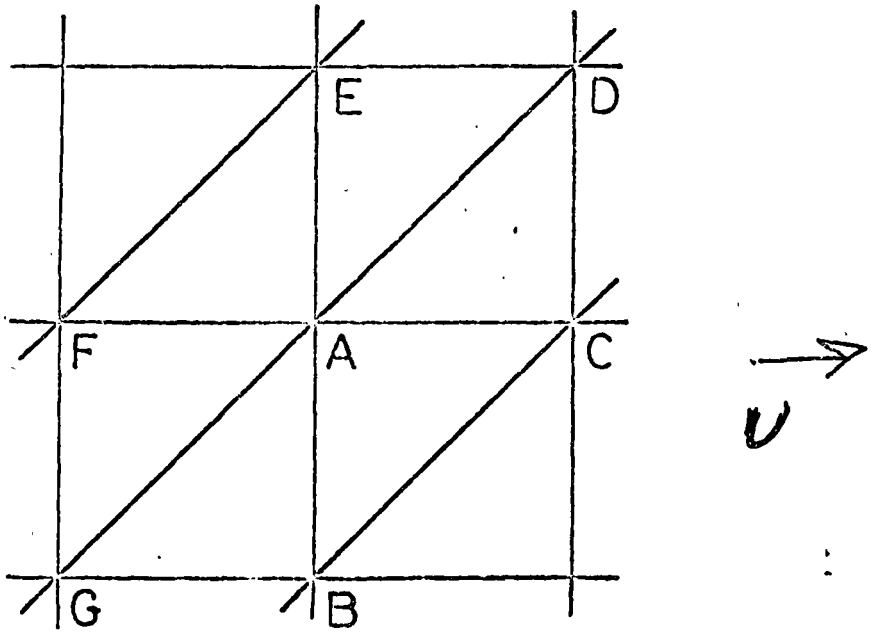


Figure 3 Regular Finite Element Grid

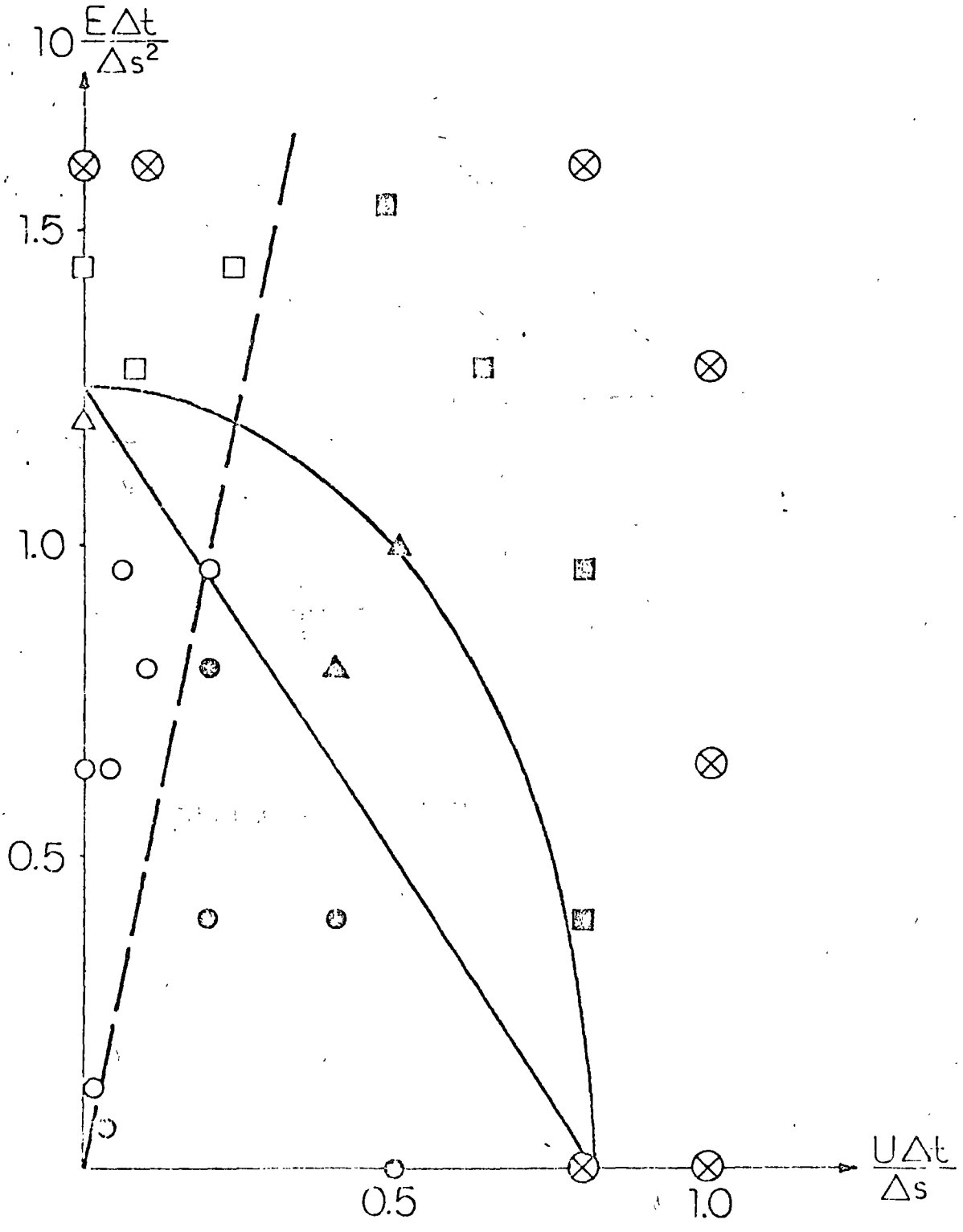


Figure 4 Comparison of 1-D Trial Runs with Theoretical Criteria

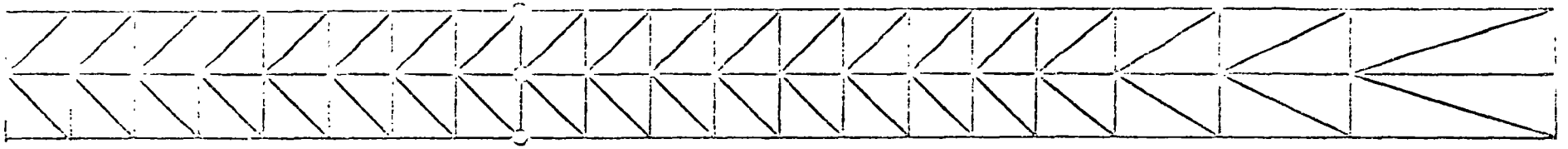


Figure 5 1-D Test Grid

Table 1

Symbols used in Figure 4

Symbol	Error after 10 Iterations	Negatives as Percent of Peak	Remarks
O	< 1%	< 10%	Good, smooth solution
O	< 1%	> 10%	Good convergence, but solution exhibits oscillations
Δ	< 10%	< 10%	Iteration error goes down rapidly with time
Δ	< 10%	> 10%	
□	> 10%	< 10%	Iteration error decreases slowly with time
□	> 10%	> 10%	
⊗	-	-	Blows up

Appendix I - REFERENCES

1. Christodoulou, G.C., et al., "Mathematical Modeling of Dispersion in Stratified Waters", M.I.T. R.M. Parsons Laboratory for Water Resources and Hydrodynamics Report No. 219, October, 1976

2. Connor, J.J. and Brebbia, C.A., "Finite Element Techniques for Fluid Flow", Butterworth, 1976

3. Gallagher, R.H., et al. (ed.), "Finite Elements in Fluids", Wiley, 1975

4. Holley, E.R., et al., "Dispersion in Homogeneous Estuary Flow", J. Hydraulics Division, ASCE, Vol. 96, HY8, August, 1970

5. Leimkuhler, W.F., et al., "Two-Dimensional Finite Element Dispersion Model", Symposium on Modeling Techniques 'Modeling 75', San Francisco, Sept., 1975

6. Price, H.S., et al., "Application of Oscillation Matrices to Diffusion-Convection Equations", J. Mathematics and Physics, Vol. 45, 1966

7. Roache, P.J., "Computational Fluid Dynamics", Hermosa Publ., Albuquerque, 1972

8. Strang, G. and Fix, G.J., "An Analysis of the Finite Element Method", Prentice-Hall, 1973

9. Taylor, C. and Davis, J.M., "Tidal Propagation and Dispersion in Estuaries", Chap. 5 in "Finite Elements in Fluids", V. 1, ed. by R.H. Gallagher et al., Wiley, 1975

10. Wellford, L.C. and Oden, J.T., "Accuracy and Convergence of Finite Element/Galerkin Approximations of Time-Dependent Problems with Emphasis on Diffusion", Chapter 2 in "Finite Elements in Fluids", V. 2, ed. by R.H. Gallagher, et al., Wiley, 1975

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION
Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS

DISPERSION IN TWO-LAYER STRATIFIED WATER BODIES

DR. JEROME J. CONNOR

MAYO, 1978.

DISPERSION IN TWO-LAYER STRATIFIED WATER BODIES

by

George C. Christodoulou[1] and Jerome J. Connor[2], M.ASCE

INTRODUCTION

During the winter season a water body is usually well-mixed through the depth. However, this is not the case in the summer. Due mainly to increased heat input near the surface, a density stratification begins to develop in the spring and by mid-summer a strong thermocline (pycnocline) often exists, dividing the water column into two distinct layers. The dynamics of such a system cannot be adequately represented by depth-averaged approximations. The effect of stratification on the flow pattern has been demonstrated by means of analytical solutions for oceans, coastal waters, and lakes, under severe simplifications of geometry and the governing equations. Ultimately, from a practical viewpoint, of main interest is not the flowfield itself, but rather the transport and dispersion of some substance in it.

To achieve a better description of both the vertical and the horizontal variability of flow in a natural water body of arbitrary geometry and bottom topography, multilayer or quasi-three-dimensional numerical models are being formulated and the development of large multi-purpose finite-difference computer codes initiated [1, 10, 17]. Although transport of constituents,

[1]Assistant Professor, Applied Hydraulics Lab., National Technical University of Athens, Athens, Greece

[2]Professor, Department of Civil Engineering, Mass. Inst. of Technology, Cambridge, MA 02139

notably water quality parameters, is being incorporated in these models, primary emphasis is placed on improving the computational techniques and software and little attention is given to such important issues as assigning values to the parameters involved, parameter sensitivity, and model verification. As the number of layers increases, model verification becomes a very difficult and costly task, as extensive field data, particularly boundary conditions, is required.

A two-layer model, while requiring minimal "tuning", provides a picture significantly different than a one-layer approach, and is quite appropriate when there is strong natural stratification. In this paper, a model is presented for the description of the dispersion of matter in such a two-layer system. After the mathematical formulation, the physical aspects of the problem are discussed, focusing primarily on quantification of the dispersion coefficients and the interfacial transport mechanisms in terms of the mean flow characteristics. The finite element method is chosen for numerical implementation because of its successful application to one-layer dispersion problems [11]. The solution procedure is discussed and its stability requirements are established. Verification of the numerical results against analytical solutions, available for simple flow conditions [6], is performed. Finally, an application of the model to the Massachusetts Bay in conjunction with a large scale field experiment serves as an example of its applicability to real world problems.

MODEL FORMULATION

The model presented herein is intended to describe the dispersion of an arbitrary constituent, possessing in general some vertical mobility, in a two-layer (coastal) water body of variable bottom topography and boundary geometry. The velocity field in both layers, as well as the layer thickness, are assumed known, presumably obtainable from a separate hydrodynamic model. By uncoupling the hydrodynamic and dispersion models, the same flow pattern can be used to investigate very economically the transport of several different substances and to experiment with various loading strategies, parameter values, etc. However, this can only be done provided the constituent of interest does not significantly affect the flow field or the density structure.

The mass balance of a constituent is expressed by the 3-D convection-diffusion equation:

$$\frac{\partial c}{\partial t} = - \frac{\partial}{\partial x} (uc + q_x) - \frac{\partial}{\partial y} (vc + q_y) - \frac{\partial}{\partial z} ((w-w_s)c+q_z) + p \qquad (1)$$

where

c is the local concentration

u,v,w are the water velocities in the x,y,z directions, respectively

w_s is the particle settling velocity, positive when in the negative z direction

q_x, q_y, q_z are diffusive fluxes

p represents generation or decay of the constituent.

Integrating (1) between the layer boundaries and using Leibnitz's rule, the equations pertaining to a layered system are obtained. Thus, for the top layer (Figure 1):

$$\frac{\partial}{\partial t}\int_{-h_1}^{\eta} c\,dz = -\frac{\partial}{\partial x}\int_{-h_1}^{\eta}(uc+q_x)dz - \frac{\partial}{\partial y}\int_{-h_1}^{\eta}(uc+q_y)dz + \int_{-h_1}^{\eta} p\,dz +$$

$$+\left[c\left(\frac{D\eta}{Dt} - w + w_s\right)-q_s\right]_{\eta} + \left[c\left(\frac{Dh_1}{Dt} + w - w_s\right)-q_i\right]_{-h_1} \qquad (2)$$

The terms in brackets represent fluxes through the layer boundaries, i.e., the free surface and interface, respectively. The kinematic condition at the surface requires

$$\left[\frac{D\eta}{Dt} - w\right]_{\eta} = 0 \qquad (3)$$

However, the interface, which is defined as the position of steepest density gradient (ideally, a density discontinuity), is not necessarily a material surface. For this surface, one can write

$$\left[w + \frac{Dh_1}{Dt}\right]_{-h_1} = w_e \qquad (4)$$

where w_e represents the relative velocity of the water particles (on the average) with respect to the layer boundary and is referred to as "entrainment" velocity. It is considered positive when upward, indicating net water motion from the bottom to the top layer. The diffusive flux component, q_i, of the interfacial transport may be expressed as a function of the concentration difference between the two layers. If the concentration at the interface is approximated by the average value of the two layers (consistent with the two-layer idealization), the overall transport from layer 2 (bottom) to

to layer 1 (top) can be written in the form:

$$Q_{21}=(w_e-w_s)c_{-h}-q_i=(w_e-w_s)\frac{c_1+c_2}{2}+\alpha(c_2-c_1) \qquad (5)$$

where α is a variable parameter. Equation (5) shows that settling counteracts entrainment, while for a neutrally buoyant constituent ($w_s=0$) the top layer would gain material through entrainment when $w_e > 0$.

With respect to the remaining terms in Equation (2), the following notation is introduced

$$C = \int_{-h_1}^{\eta} c\,dz = \bar{c}_1 H_1$$

$$c = \bar{c} + c'' \,, \quad u = \bar{u} + u'' \,, \quad v = \bar{v} + v'' \qquad (6)$$

where the overbar denotes the average value over the layer thickness and the double prime represents spatial deviation from the average. Equation (2) now takes the form:

$$\frac{\partial C_1}{\partial t} + \frac{\partial(\bar{u}_1 C_1)}{\partial x} + \frac{\partial(\bar{v}_1 C_1)}{\partial y} = -\frac{\partial}{\partial x} Q_{x_1} - \frac{\partial}{\partial y} Q_{y_1} + P_1 \qquad (7)$$

Sources, decay, and boundary flux terms are included in P_1. The total dispersive fluxes, Q, account for both horizontal turbulent diffusion and dispersion due to vertical velocity nonuniformities and are assumed to be approximated by Fickian expressions:

$$Q_{x_1} = \int_{-h_1}^{\eta} (u''c''+q_x)\,dz = -H_1\left(E_{xx_1}\frac{\partial \bar{c}_1}{\partial x} + E_{xy_1}\frac{\partial \bar{c}_1}{\partial y}\right)$$

$$Q_{y_1} = \int_{-h_1}^{\eta} (v''c''+q_y)\,dz = -H_1\left(E_{yx_1}\frac{\partial \bar{c}_1}{\partial x} + E_{yy_1}\frac{\partial \bar{c}_1}{\partial y}\right) \qquad (8)$$

The overall dispersion coefficients are elements of a second order tensor, consisting of an eddy diffusivity component and a shear dispersion component:

$$\underset{\sim}{E} = \begin{bmatrix} E_{xx} & E_{xy} \\ E_{yx} & E_{yy} \end{bmatrix} = \underset{\sim}{\varepsilon} + \underset{\sim}{E}^{d} \qquad (9)$$

Their quantification is discussed in the next section.

Following the same approach, one obtains the integrated equations for the bottom layer:

$$\frac{\partial C_2}{\partial t} + \frac{\partial (\bar{u}_2 C_2)}{\partial x} + \frac{\partial (\bar{v}_2 C_2)}{\partial y} = - \frac{\partial}{\partial x} Q_{x_2} - \frac{\partial}{\partial y} Q_{y_2} + P_2 \qquad (10)$$

where P_2 will, in general, include deposition to the bottom. Equations (7) and (10) are the governing equations of the two-layer system.

The boundary conditions for the dispersion problem are of two types (see Figure 2):

(i) concentration specified: $C = C^*$ on S_c

(ii) normal dispersive flux (i.e., concentration gradient) specified: $Q_n = Q_n^*$ on S_q.

Assuming reflecting land boundaries, one usually prescribes $Q_n = 0$. Of major concern is the treatment of the ocean (open) boundary. The concentration may be maintained at zero only as long as the plume remains sufficiently far from the boundary. An ideal, but not economical, solution is to make the grid so large as to ensure that the plume will always remain well within the computational domain.

In practice, different conditions are prescribed for outflow and inflow boundary segments. In the former case, the concentration gradient is commonly specified, its value obtained by extrapolation from the interior. In the latter case, the concentration should be specified, but this is difficult since it is related to mixing conditions outside of the domain being modeled [10]. A simple procedure, which proves satisfactory when the plume reaches the boundary with a low concentration gradient at a segment of predominantely outward flow, is to specify and maintain the gradient at zero [16]. This allows the material to advect through the boundary and assumes essentially complete mixing in its neighborhood.

DISPERSION COEFFICIENTS

The horizontal spreading of a constituent within a given "layer" is accomplished by the following three mechanisms:

(a) Advection, in particular spatial or temporal variability of layer average currents,

(b) Turbulent diffusion, i.e., mixing due to small or large scale turbulent eddies, and

(c) Dispersion due to vertical shear, that is, velocity nonuniformities over the layer thickness.

The contribution of the last two mechanisms is commonly expressed by the introduction of diffusion and dispersion coefficients, by analogy to the molecular mixing process.

These coefficients basically arise from the simplified representation of the velocity field, and, indeed, the more simplified the latter becomes, the larger the coefficients need to be. Considering the diffusion of a cloud due to turbulence, it is noticed that, at first, moderate size eddies contribute to the advection of the cloud as a whole, while mixing takes place at very small scales. As the size of the cloud increases with time, larger eddies become involved in its internal mixing. Thus, it is found that diffusion coefficients increase with time (or cloud size) [13]. However, when the flow field is specified at a certain spatial discretization, such continuous growth of the diffusion coefficients is not justified once the cloud increases beyond the level of discretization, since eddies of the scale of the grid size are still described by the advection terms.

One approach for quantifying the eddy diffusivity is based on using the 4/3 diffusion law derived from the theory of locally isotropic turbulence. This is applicable to horizontal diffusion in the ocean if one assumes that the eddies are essentially isotropic horizontally [14]. Then

$$\varepsilon = be^{1/3}L^{4/3} \qquad (10)$$

where

ε is the eddy diffusion coefficient

e is the rate of energy input, per unit surface area

L is the length scale, presumably related to the grid size

b is a numerical constant, or order 0.1.

An alternative expression for ε, based on mixing length arguments [2,8] has the advantage of using the readily available mean velocity gradients:

$$\varepsilon = L^2\sqrt{\phi} \tag{11}$$

where

$$\phi = 2\left(\frac{\partial \bar{u}}{\partial x}\right)^2 + 2\left(\frac{\partial \bar{v}}{\partial y}\right)^2 + \left(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x}\right)^2$$

According to [8], the sub-grid scale eddy coefficient is modeled using a length scale which is an order or magnitude smaller than the grid size. However, the resolution of the flow field description, associated with spatial averaging in the hydrodynamic model being employed, has to be taken into account and the coefficient increased accordingly.

The analogy between the effective horizontal spreading due to nonuniformity of the velocity profile and the turbulent diffusion process was shown initially by Taylor [18] for steady 1-D flow through a pipe and later by Elder [9] for open channel flow. A parallel argument for the case of two horizontal velocity components (u,v) shows that the representation of the dispersive fluxes according to (8) is indeed appropriate [6]. The shear dispersion coefficients are identified as:

$$E^d_{xx} = \frac{1}{H}\int_0^H \frac{1}{\varepsilon_z}\left[\int_0^z u''d\zeta\right]^2 dz$$

$$E^d_{yy} = \frac{1}{H}\int_0^H \frac{1}{\varepsilon_z}\left[\int_0^z v''d\zeta\right]^2 dz \tag{12}$$

$$E^d_{xy} = E^d_{yx} = \frac{1}{H}\int_0^H \frac{1}{\varepsilon_z}\left[\int_0^z u''d\zeta\right]\left[\int_0^z v''d\zeta\right]dz$$

where ε_z is the vertical diffusion coefficient, and H is the layer thickness. Once the velocity profile is specified, the above integrals can be evaluated. One can prove that the shear dispersion coefficients comprise a second order tensor [6]. Therefore, profile information is needed only along any two perpendicular directions. In the simplest case, if self-similarity of the velocity profile in a layer is assumed, the dispersion coefficients may be related directly to the mean velocity and layer thickness.

This treatment of shear dispersion is valid, only when some "initial time" has passed after the introduction of the material. The initial time is related to the vertical mixing time scale, given by [3]:

$$T_c = H^2/\pi^2\varepsilon_z \tag{13}$$

Furthermore, with respect to the effect of a tidal component, it has been found [3] that the dispersion coefficient is essentially the same as if the flow was steady at any point of the tidal cycle provided that

$$T/T_c \gtrsim 1 \tag{14}$$

For typical values, H ≃ 10-20 m (30-60 ft), ε_z = 50 cm^2/sec (0.05 ft^2/sec), one obtains T_c ≃ 0.5 - 2 hours, and condition (14) holds.

INTERFACIAL TRANSPORT

The erosion of the quiescent lower layer by the upper layer, moving under the influence of wind or other driving mechanism, is a well-known phenomenon in water bodies. It is explained by the one-way transport from the nonturbulent to the turbulent layer, often called entrainment. The term is most familiar in the context of a jet, which draws ambient fluid due to its high momentum. The mechanism of erosion of a distinguishable density interface has been examined in the past, mostly experimentally. Turbulent eddies appear to scour the interface, sweeping away interfacial disturbances at relatively large time intervals [19]. When both layers are turbulent and have comparable velocities, as is usually the case in coastal waters, there must be a two-way transfer. Denoting the respective volumetric rates of transport, per unit area, by m_{21} and m_{12} (see Figure 3), the net rate of transport of material toward the top layer is

$$Q_{21} = m_{21}c_2 - m_{12}c_1 \tag{15}$$

This, of course, assumes that the particles of interest do not have independent motion, which would provide yet another contribution to interfacial transport. By setting $w_s = 0$ in Equation (5) and comparing with Equation (15), it is evident that

$$\begin{aligned} w_e &= m_{21} - m_{12} \\ \alpha &= (m_{12} + m_{21})/2 \end{aligned} \tag{16}$$

Thus, the entrainment velocity expresses the net rate of water motion between the layers, while α represents an average exchange rate and indicates that interfacial transport is present even without net entrainment (i.e., "steady-state" conditions with respect to layer boundaries).

Several experimental and theoretical investigations [4, 12, 19] have been carried out in the past, mostly in 1-D two-layer systems with one layer quiescent. The one-way transport rate (similar to m_{ij}, above) was then determined by the thickening of the moving layer. It was found proportional to some characteristic velocity and inversely proportional to a Richardson number associated with the stability of the system. Various length and velocity scales have been used and, at first, the agreement between a number of the proposed formulas seems to be only qualitative. However, if comparable measures are used, a rather general expression emerges in the form [6]:

$$m_{ji} \simeq \frac{10^{-3}|V_i|}{Ri_o} \qquad i,j = 1,2 \quad (i \neq j) \tag{17}$$

where, in a 2-D domain, the overall Richardson number is defined as

$$Ri_o = \frac{g \frac{\Delta\rho}{\rho} \bar{H}}{(\vec{V}_i - \vec{V}_j)^2} \tag{18}$$

and $\Delta\rho$ is the (small) density difference

ρ is the density of either layer

$\bar{H}$ is the average layer thickness, i.e., half the depth

V_i, V_j are the depth-average velocities in layers, i,j.

The rates of interfacial transport are typically small for stably stratified water bodies; e.g., the value of α is of the order of 10^{-5} m/sec (ft/sec). Nevertheless, this is of the order of settling velocities of fine particles and its contribution may become significant over the relatively large length and time scales typical of coastal areas - especially when multiplied by a large concentration difference.

The use of Equation (17) is conditional on the existence of a mean velocity in the layer, which is usually the case in tidally dominated flows in coastal areas. However, interfacial transport may well be present in the absence of mean flow, as indicated by experiments with stirring grids [19]. Further research is needed in this area for a more general quantification.

NUMERICAL TECHNIQUE

The finite element method is chosen for spatial discretization because of its great flexibility in grid layout and easy handling of spatial variability. To find the approximate solution the weighted residual method is applied to each layer, resulting in the symmetrical weak form [11]. Then, the

domain is subdivided into linear triangular elements, resulting in a set of linear ordinary differential equations with the nodal concentrations as unknowns;

$$\begin{aligned} \underset{\sim}{M}\ \dot{\underset{\sim}{C}}_1 &= \hat{\underset{\sim}{P}}_1 \\ \underset{\sim}{M}\ \dot{\underset{\sim}{C}}_2 &= \hat{\underset{\sim}{P}}_2 \end{aligned} \tag{19}$$

In Equation (19), $\underset{\sim}{M}$ is the geometrical system matrix (it is assumed that the spatial discretization is the same for both layers), and $\hat{\underset{\sim}{P}}_i$ is the forcing vector for layer i containing advection, dispersion, decay, interfacial transfer, sources and boundary conditions. Time integration is performed by an implicit iterative trapezoidal scheme, as follows:

$$\begin{aligned} \underset{\sim}{C}_{1,t+\Delta t}^{(i+1)} &= \underset{\sim}{C}_{1,t} + \frac{\Delta t}{2}\ \underset{\sim}{M}^{-1}\ (\hat{\underset{\sim}{P}}_{1,\,t+\Delta t}^{(i)} + \hat{\underset{\sim}{P}}_{1,t}) \\ \underset{\sim}{C}_{2,t+\Delta t}^{(i+1)} &= \underset{\sim}{C}_{2,t} + \frac{\Delta t}{2}\ \underset{\sim}{M}^{-1}\ (\hat{\underset{\sim}{P}}_{2,t+\Delta t}^{(i)} + \hat{\underset{\sim}{P}}_{2,t}) \end{aligned} \tag{20}$$

Since $\underset{\sim}{M}$ is time invariant, it has to be inverted only once. By lumping all other terms in $\hat{\underset{\sim}{P}}_i$, maximum generality in handling time variability or nonlinearity of the relevant parameters and loadings is achieved. In practice, the iteration continues until either a measure of the difference between consecutive iterates is below a specified tolerance, or the number of iterations reaches an imposed upper limit.

In the case of constant and equal layer thicknesses, no entrainment and constant α, Equation (19) can be written in the expanded form [6]:

$$\underset{\sim}{M}\,\dot{\underset{\sim}{C}}_1 + (\underset{\sim}{A}_1 + \underset{\sim}{K}_1 + \underset{\sim}{D}_1)\,\underset{\sim}{C}_1 + \underset{\sim}{G}(\underset{\sim}{C}_1 - \underset{\sim}{C}_2) = \underset{\sim}{S}_1 - \underset{\sim}{F}_1^b$$

$$\underset{\sim}{M}\,\dot{\underset{\sim}{C}}_2 + (\underset{\sim}{A}_2 + \underset{\sim}{K}_2 + \underset{\sim}{D}_2)\,\underset{\sim}{C}_2 + \underset{\sim}{G}(\underset{\sim}{C}_2 - \underset{\sim}{C}_1) = \underset{\sim}{S}_2 - \underset{\sim}{F}_2^b \qquad (21)$$

where $\underset{\sim}{A}_i$, $\underset{\sim}{K}_i$, $\underset{\sim}{D}_i$, $\underset{\sim}{G}_i$ are the advection, dispersion, decay, and interfacial diffusion matrices; $\underset{\sim}{S}_i$ contains source loadings; and $\underset{\sim}{F}_i^b$ includes the terms resulting from the prescribed boundary conditions.

The trapezoidal integration scheme applied to Equations (21) can be shown to be unconditionally stable, under no iteration, for an arbitrary grid [6]. The procedure is similar to that of the one layer case [7]. Actually, the inter-layer exchange term enhances the stability of the system. However, the iterative procedure used imposes a restriction on the time step.

Applying the time integration scheme to (21) yields:

$$\underset{\sim}{M}\,\underset{\sim}{C}_{1,n+1}^{(i+1)} = -\frac{\Delta t}{2}\,(\underset{\sim}{A}_1+\underset{\sim}{K}_1+\underset{\sim}{D}_1+\underset{\sim}{G})_{n+1}\,\underset{\sim}{C}_{1,n+1}^{(i)} + \frac{\Delta t}{2}\,\underset{\sim}{G}_{n+1}\,\underset{\sim}{C}_{2,n+1}^{(i)} + \underset{\sim}{Q}_1$$

$$\underset{\sim}{M}\,\underset{\sim}{C}_{2,n+1}^{(i+1)} = -\frac{\Delta t}{2}(\underset{\sim}{A}_2+\underset{\sim}{K}_2+\underset{\sim}{D}_2+\underset{\sim}{G})_{n+1}\,\underset{\sim}{C}_{2,n+1}^{(i)} + \frac{\Delta t}{2}\,\underset{\sim}{G}_{n+1}\,\underset{\sim}{C}_{1,n+1}^{(i+1)} + \underset{\sim}{Q}_2 \qquad (22)$$

where the subscript n+1 refers to time $t+\Delta t$ and the quantities $\underset{\sim}{Q}_1$, $\underset{\sim}{Q}_2$ are known from the previous time step. To investigate convergence of the iteration procedure, the equations are written as:

$$\underset{\sim}{R}\,\underset{\sim}{C}^{(i+1)} = \underset{\sim}{B}\,\underset{\sim}{C}^{(i)} + \underset{\sim}{Q} \qquad (23)$$

where

$$\underset{\sim}{C}^{(i+1)} = \begin{bmatrix} \underset{\sim}{C}_{1,n+1}^{(i+1)} \\ \underset{\sim}{C}_{2,n+1}^{(i+1)} \end{bmatrix} \qquad \underset{\sim}{C}^{(i)} = \begin{bmatrix} \underset{\sim}{C}_{i,n+1}^{(i)} \\ \underset{\sim}{C}_{2,n+1}^{(i)} \end{bmatrix} \qquad \underset{\sim}{R} = \begin{bmatrix} \underset{\sim}{M} & \underset{\sim}{0} \\ -\frac{\Delta t}{2}\underset{\sim}{G}_{n+1} & \underset{\sim}{M} \end{bmatrix}$$

$$\underset{\sim}{B} = \begin{bmatrix} -\frac{\Delta t}{2}(\underset{\sim}{A}_1+\underset{\sim}{K}_1+\underset{\sim}{D}_1+\underset{\sim}{G})_{n+1} & \frac{\Delta t}{2}\underset{\sim}{G}_{n+1} \\ \underset{\sim}{0} & -\frac{\Delta t}{2}(\underset{\sim}{A}_2+\underset{\sim}{K}_2+\underset{\sim}{D}_2+\underset{\sim}{G})_{n+1} \end{bmatrix}$$

The convergence requirement is

$$||\underset{\sim}{R}^{-1}\ \underset{\sim}{B}|| < 1 \tag{24}$$

Expressing $\underset{\sim}{R}$ as a product,

$$\underset{\sim}{R} = \begin{bmatrix} \underset{\sim}{M} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{M} \end{bmatrix} \begin{bmatrix} \underset{\sim}{I} & \underset{\sim}{0} \\ -\frac{\Delta t}{2}\underset{\sim}{M}^{-1}\underset{\sim}{G}_{n+1} & \underset{\sim}{I} \end{bmatrix} \tag{25}$$

leads to

$$\underset{\sim}{R}^{-1} = \begin{bmatrix} \underset{\sim}{I} & \underset{\sim}{0} \\ \frac{\Delta t}{2}\underset{\sim}{M}^{-1}\underset{\sim}{G}_{n+1} & \underset{\sim}{I} \end{bmatrix} \begin{bmatrix} \underset{\sim}{M}^{-1} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{M}^{-1} \end{bmatrix} \tag{26}$$

Since now both $\underset{\sim}{R}^{-1}$ and $\underset{\sim}{B}$ involve triangular matrices, their eigenvalue norms are conveniently expressed in terms of their diagonal elements, Thus, condition (24) is equivalent to

$$||\underset{\sim}{M}^{-1}\ \frac{\Delta t}{2}(\underset{\sim}{A}_i+\underset{\sim}{K}_i+\underset{\sim}{D}_i+\underset{\sim}{G})_{n+1}|| < 1 \qquad i = 1,2 \tag{27}$$

which implies

$$\Delta t < \frac{2}{||\underset{\sim}{M}^{-1}(\underset{\sim}{A}_i+\underset{\sim}{K}_i+\underset{\sim}{D}_i+\underset{\sim}{G}||} \tag{28}$$

A more restrictive condition is

$$\Delta t < \frac{2}{||\underset{\sim}{M}^{-1}\underset{\sim}{A}_i||+||\underset{\sim}{M}^{-1}\underset{\sim}{K}_i||+||\underset{\sim}{M}^{-1}\underset{\sim}{D}_i||+||\underset{\sim}{M}^{-1}\underset{\sim}{G}||} \tag{29}$$

These criteria are analogous to the one-layer results [7], the difference being the addition of the interfacial exchange term.

Evaluating the matrix expressions for an individual equilateral triangular element of side Δs yields an approximation to (29) in terms of the problem parameters:

$$\Delta t < \frac{1}{1.2\,\frac{V_i}{\Delta s} + 8\,\frac{E_i}{\Delta s^2} + \frac{k}{2} + \frac{\alpha'}{2}} \tag{30}$$

where k is the decay rate, $\alpha' = \alpha/\bar{H}$, E_i is the (assumed isotropic) dispersion coefficient and V_i is the (assumed uniform) velocity of layer i. As discussed earlier, the value of α' is commonly small and its contribution to limiting the time step will typically be marginal. Then, Δt is basically restricted by the flow conditions in the individual layers.

VERIFICATION

To test the accuracy of the numerical approximation, comparisons with analytical solutions are desirable. However, the latter are available only under very simple flow conditions.

In Figure 5, the numerical model is compared with the solution for an instantaneous source in the top layer of a 1-D counterflow, derived in [6]. A unit load is distributed between the three nodes at x = 0 of the grid whown in Figure 4, and the results adjusted to yield values per unit width of the channel. A unit depth is assumed for each layer. Zero concentration is specified at the ends of the grid and zero flux is prescribed along the side boundaries. The parameters used are:

$V_1 = -V_2 = 0.05$ m/sec (0.164 ft/sec)

$E_1 = E_2 = 0.01$ m^2/sec (0.108 ft^2/sec)

$\alpha = 5 \times 10^{-4}$ m/sec (16.4×10^{-3} ft/sec) $\quad k = 0$

$\Delta t = 0.1$ sec

Very good agreement with the analytical solution is obtained. The much lower concentrations observed in the bottom layer support to some extent the traditional treatment of the interface as a barrier. However, this simplification may not be reasonable for long time periods and is certainly not valid for substances possessing vertical mobility. Figure 5 also points out the great advantage of the two-layer treatment, in relation to the more detailed description of the flow field. In this particular counterflow case, the depth-average velocity is zero and a one-layer approach would imply a stationary concentration peak located at the origin.

The behavior of the model at steady state was also examined. The results for a continuous load, of one unit/sec, introduced in the top layer, are shown in Figure 6. A high decay rate is

specified to speed up the arrival at steady-state and keep appreciable concentrations away from the boundary. The remaining parameters are the same as in the transient test, except for a higher interfacial mixing rate used to make the exchange between the layers more pronounced. Again agreement is quite good.

APPLICATION

To establish confidence in the predictive capability of the model and the degree of its applicability under natural conditions, further verification consisting of comparison to real world cases is necessary. Agreement can never be expected to be perfect, in view of the extreme complexity of the physical processes involved and the unavoidable simplifications employed in any model. Nevertheless, the ability of the model to reproduce certain basic features of the actual data should be evaluated.

A dispersion experiment was carried out by the R.M. Parsons Laboratory of M.I.T., sponsored by the Boston Edison Co., in the vicinity of the Pilgrim Nuclear Power Station on the Massachusetts Coast (Figure 7), in August 1975. Five hundred pounds of small sphalerite particles (ZnS) with fluorescent inclusions were introduced into the water and their motion was subsequently monitored for five days through samples taken by boat and by helicopter. By averaging, at each location, samples taken above and below the thermocline, the field data

were reduced to a single representative value for each layer. The resulting plots, in particles/lt, are shown in Figures 8a, b, c, corresponding to 1, 2 and 3 days after the dumping took place. The plume is seen to move slowly to the southeast, approximately parallel to the shoreline and later extend to the east.

In the numerical simulations, the finite element grid was the same as used in previous applications of one-layer models to the Bay [15, 16]. The shaded triangle was loaded over a period of three timesteps (i.e., 4500 sec.), which corresponds approximately to the actual duration of the dumping. However, the area of the triangle is quite large in comparison to the actual source and consequently one should expect unrealistically large plume areas for short times. The value of the (isotropic) dispersion coefficient, 30 m^2/sec,(323 ft^2/sec), and the difference in tidal amplitude between the ends of the open boundary were kept the same as established for the one layer models [6]. The circulation model used to provide velocity inputs is that of Wang and Connor [20]. Since this requires, at present, that both layers extend over the whole domain, some nodal depths had to be artificially increased to at least 15 m (49.2 ft), in order to avoid intersection of the interface with the bottom. As initial condition, the position of the interface was set at 8 m (26.2 ft), consistent with the little available information [5]. Along the ocean boundary the interface was assumed to vary linearly and move together with the free surface over the tidal cycle.

The sensitivity to the type of interface motion was found small [6], but other than linear configurations were not examined. Actual time varying wind data were used in the computations, while a "typical" tidal cycle was used repeatedly. The interfacial mixing rate was set at 10^{-5} m/sec (3.28×10^{-5} ft/sec) and the settling velocity at 7.3×10^{-5} m/sec (24×10^{-5} ft/sec), based on an average particle size of 7 microns.

Computed concentrations at 1, 2, 3 days after the injection are shown in Figures 9a, b, c. Taking into account the initial spreading of the source and the uncertainties about the velocity field, good qualitative agreement is observed, with respect to the location and peak values of the plume.

CONCLUSIONS

In this paper, the problem of dispersion in strongly stratified water bodies is examined. The two-layer idealization is adopted as a useful extreme case and, at the same time, the simplest to handle mathematically. Quantitative expressions for the dispersion coefficients and the interfacial transport rates, needed for engineering applications, are proposed. Also, a criterion for selecting the time step is presented.

The ability of the two-layer model to handle transport between the layers was seen to be important in providing a refined picture of the vertical concentration distribution, whether or not the constituent of interest has some vertical mobility. A further advantage of the two-layer formulation

lies in the more detailed description of the velocity field. This, in turn, points out the need for using realistic current inputs.

The development of numerical techniques is outgrowing the present ability to define realistic inputs and also the basic knowledge of some of the physical processes involved. Further fundamental research is needed for better understanding the turbulent mixing process in stratified environments. Also, field monitoring programs are required to provide reliable inputs, primarily on the behavior of the interface along open boundaries.

ACKNOWLEDGEMENTS

The research leading to this paper was performed while the first of the authors was a graduate student at the Massachusetts Institute of Technology. Financial support was obtained from the Boston Edison Co., and the Sea Grant Office of the National Oceanic and Atmospheric Administration.

REFERENCES

1. Abbott, M.B., et al., "Systems Modeling in Stratified Fluids", Symposium on Modeling Techniques 'Modeling 75', San Francisco, Sept., 1975.

2. Boericke, P.R. and Hall, D.W., "Hydraulic and Thermal Dispersion in an Irregular Estuary", J. Hydraulics Div., ASCE, Vol. 100, HY 1, Jan., 1974

3. Bowden, K.F., "Fundamentals of Dispersion", Proc. of Symp. on Mathematical and Hydraulic Modeling of Estuarine Pollution, at the Water Pollution Research Labs, April, 1972

4. Breusers, H.N.C. (ed.), "Momentum and Mass Transfer in Stratified Flows", Literature Study, Delft Hydraulics Lab. Report R880, Dec., 1974

5. Bumpus, D.F., "Review of the Physical Oceanography of Massachusetts Bay", Woods Hole Oceanographic Inst. Report 74-8, Feb., 1974

6. Christodoulou, G.C., et al., "Mathematical Modeling of Dispersion in Stratified Waters", R.M. Parsons Laboratory for Water Resources and Hydrodynamics, Tech. Report No. 219, M.I.T., Oct., 1976

7. Christodoulou, G.C. and Connor, J.J., "Finite Element Stability for Convection-Diffusion", J. Hydraulics Div., ASCE, submitted for publication

8. Deardorff, J.W., "On the Magnitude of the Subgrid Scale Eddy Coefficient", J. Computational Physics, Vol. 7, Part 1, 1971

9. Elder, J.W., "The Dispersion of Marked Fluid in Turbulent Shear Flow", J. Fluid Mechanics, Vol. 5, Part 4, 1959

10. Leendertse, J.J. and Liu, S.K., "A Three-Dimensional Model for Estuaries and Coastal Seas: Vol. II, Aspects of Computation", Rand Co., June, 1975

11. Leimkuhler, W.F., et al., "Two-Dimensional Finite Element Dispersion Model", Symp. on Modeling Techniques 'Modeling 75', San Francisco, Sept., 1975

12. Long, R.R., "The Influence of Shear on Mixing Across Density Interfaces", J. Fluid Mechanics, Vol. 70, Part 2, July 1975

13. Okubo, A., "Oceanic Diffusion Diagrams", Deep Sea Research, Aug., 1971

14. Osmidov, R.V., "On the Turbulent Exchange in a Stably Stratified Ocean", Izv. Acad. of Science USSR, Atmospheric and Oceanic Physics, Vol. 1, No. 8, 1965

15. Pagenkopf, J.R., et al, "Circulation and Dispersion Studies at the Pilgrim Nuclear Power Station, Rocky Point, Mass", R.M. Parsons Laboratory for Water Resources and Hydrodynamics, Technical Report 210, M.I.T., Feb., 1976

16. Pearce, B.R. and Christodoulou, G.C., "Application of a Finite Element Dispersion Model for Coastal Waters", Proc. of XVI I.A.H.R. Congress, Sao Paulo, 1975

17. Simons, T.J., "Development of Three-Dimensional Numerical Models of the Great Lakes", Scientific Series No. 12, Canada Centre for Inland Waters, Burlington, Ontario, 1973

18. Taylor, G.I., "The Dispersion of Matter in Turbulent Flow Through a Pipe", Proc. Royal Soc. of London, Ser. A, V. 223, 1954

19. Turner, J.S., "Buoyancy Effects in Fluids", Cambridge Univ. Press, 1973

20. Wang, J.D. and Connor, J.J., "Finite Element Model of Two Layer Coastal Circulation", Proc. of XIV Inter. Conf. on Coastal Engrg., Copenhagen, 1974

NOTATION

b	numerical constant
c	local concentration
$\bar{c}_1$, $\bar{c}_2$	layer average concentrations
$c"$	concentration deviation from average value
e	rate of energy influx
$-h_1$	interface elevation
$-h_2$	bottom elevation
(i)	iteration index
k	decay rate
m_{12}, m_{21}	one-way interfacial transport rates
n	time discretization index
p	internal volumetric source/sink term
q_x, q_y, $\bar{q}_z$	turbulent diffusion fluxes
q_i, q_s	diffusive fluxes through interface and surface
t	time
u, v, w	local velocity components
$\bar{u}$, $\bar{v}$	layer average velocities
$u"$, $v"$	velocity deviations from average values
w_e	entrainment velocity
w_s	settling velocity
x, y, z	cartesian coordinates
$\underset{\sim}{A}$	advection matrix
$\underset{\sim}{B}$, $\underset{\sim}{Q}$, $\underset{\sim}{R}$	matrices
C_1, C_2	layer-integrated concentrations

$\underset{\sim}{D}$	decay matrix
E, E_{xx}, E_{yy}, E_{xy}	overall dispersion coefficients
E^d, E^d_{xx}, E^d_{yy}, E^d_{xy}	shear dispersion coefficients
$\underset{\sim}{F}^b$	boundary conditions vector
$\underset{\sim}{G}$	interfacial transport matrix
H_1, H_2	layer depths
$\bar{H}$	average of H_1 and H_2
$\underset{\sim}{K}$	dispersion matrix
L	length scale for diffusion
$\underset{\sim}{M}$	geometrical matrix
P	term including sources/sinks, b.c. and interfacial transfer
$\underset{\sim}{\hat{P}}$	overall forcing vector
Q_x, Q_y, Q_n	dispersive fluxes
Q_{21}	overall interfacial transport
R_{io}	overall Richardson number
$\underset{\sim}{S}$	source vector
S_c, S_q	boundary segment where concentration, or dispersive flux, is specified
T	tidal period
T_c	time scale for mixing
V_1, V_2	layer velocities
α	proportionality factor for interfacial diffusion
α'	$\alpha/\bar{H}$
ε, ε_x, ε_y	turbulent diffusion coefficients
η	surface elevation
ρ, $\Delta\rho$	density, density difference
Δs	grid size
Δt	time step

FIGURE CAPTIONS

FIGURE 1 The Two-Layer Idealization

FIGURE 2 Solution Field and Boundary Conditions

FIGURE 3 Schematization of Interfacial Transport

FIGURE 4 One-Dimensional Finite Element Test Grid

FIGURE 5 1-D Distribution at t = 10 sec After an Instantaneous Injection

FIGURE 6 1-D Steady State Distribution for a Continuous Load

FIGURE 7 Massachusetts Bay Finite Element Grid

FIGURE 8
(a) Experimental Results at Day D+1
(b) Experimental Results at Day D+2
(c) Experimental Results at Day D+3

FIGURE 9
(a) Computed Concentrations at Day D+1
(b) Computed Concentrations at Day D+2
(c) Computed Concentrations at Day D+3

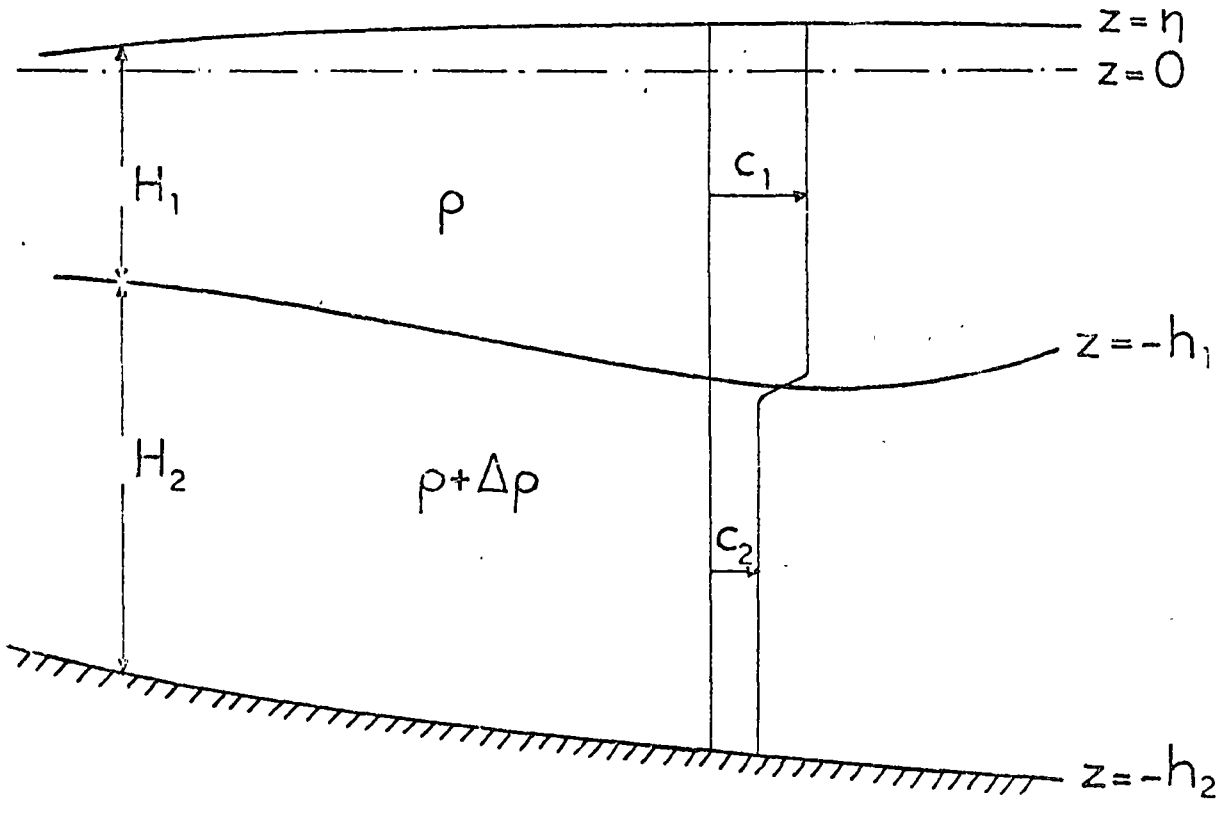


Figure 1 The Two-Layer Idealization

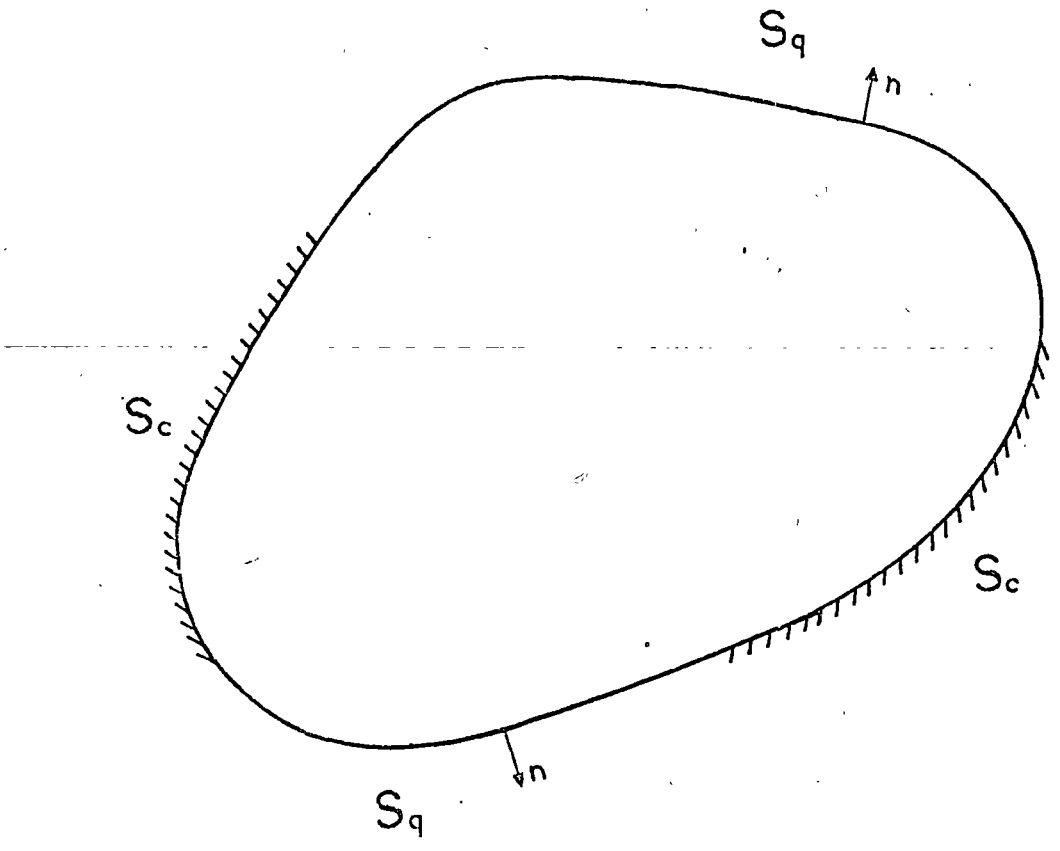


Figure 2 Solution Field and Boundary Conditions

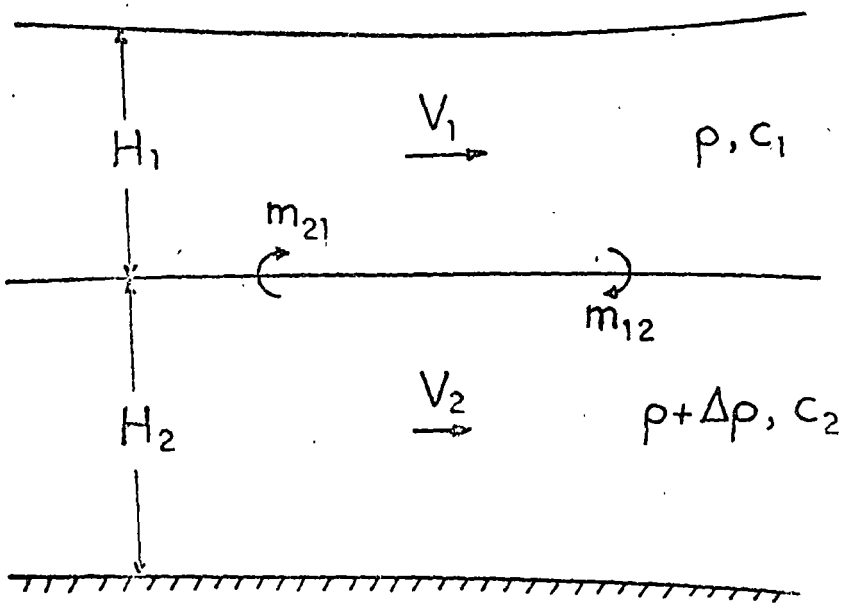


Figure 3 Schematization of Interfacial Transport

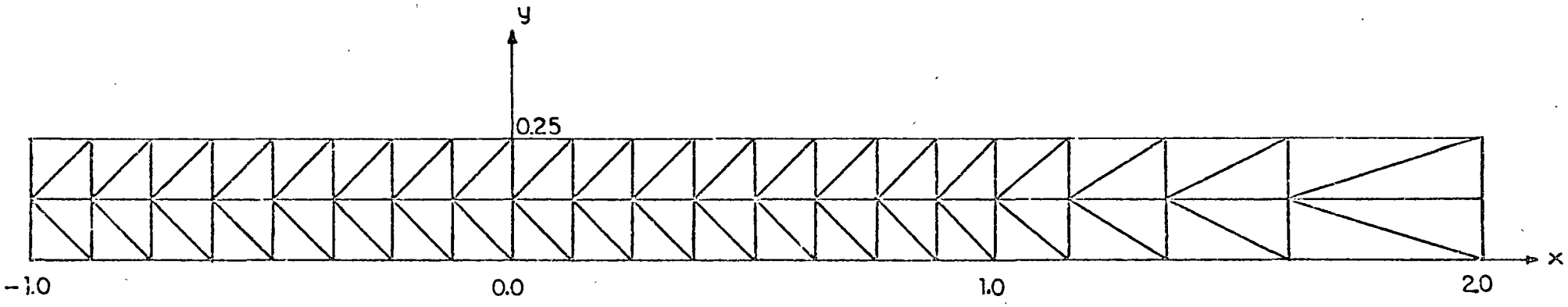


Figure 4 One-Dimensional Finite Element Test Grid

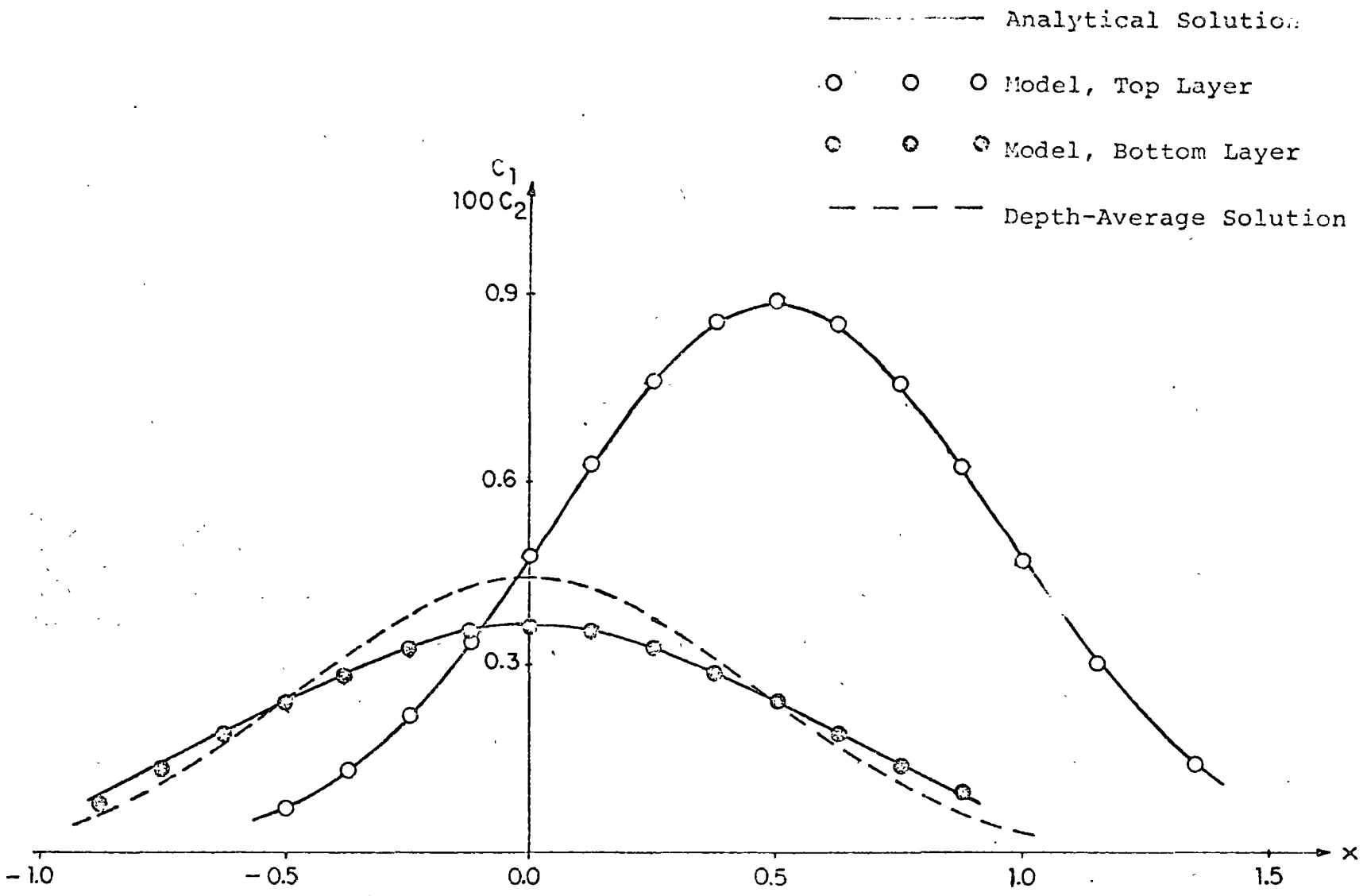


Figure 5 1-D Distribution at t=10 sec. after an Instantaneous Injection

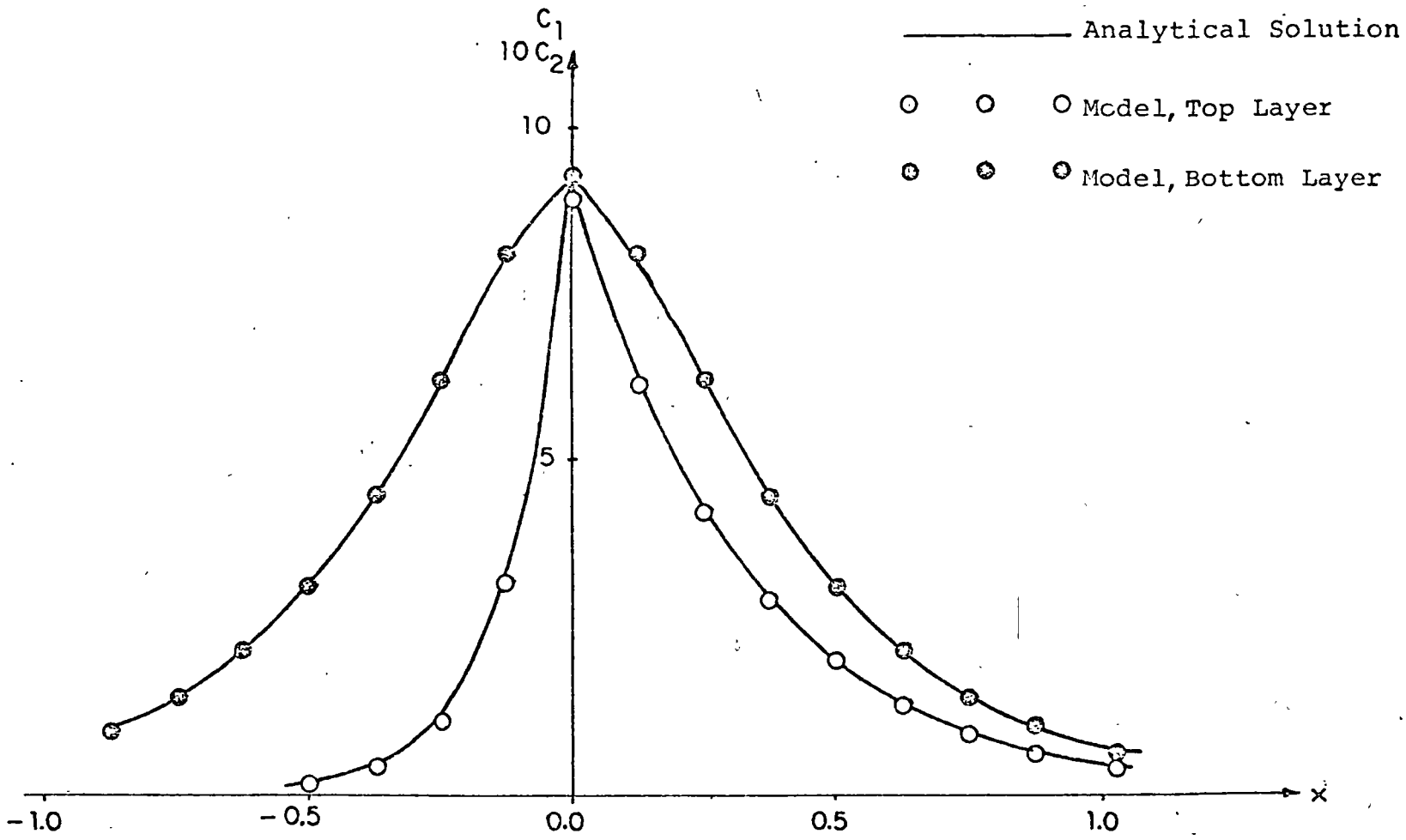


Figure 6 1-D Steady State Distribution for a Continuous Load

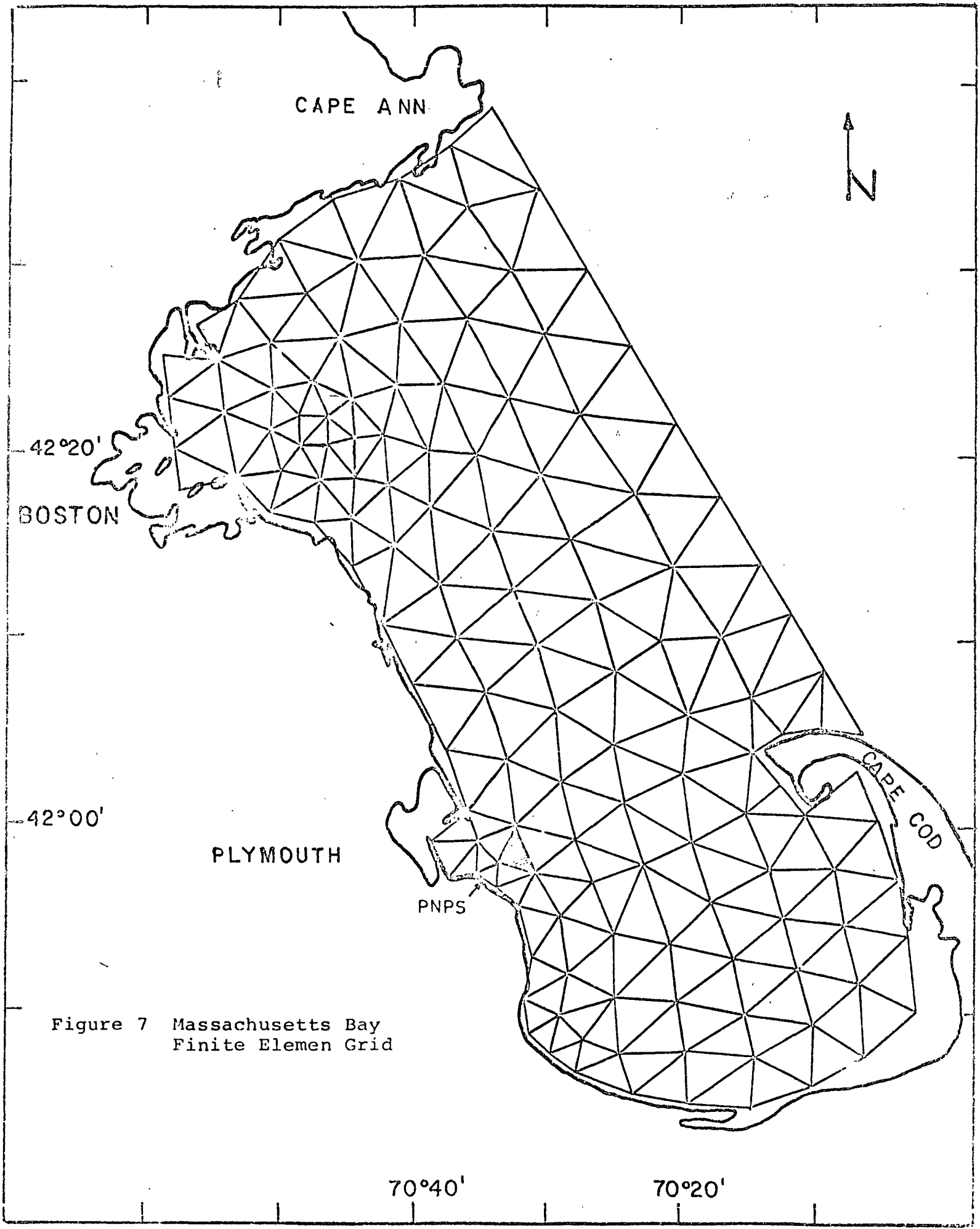


Figure 7 Massachusetts Bay
Finite Elemen Grid

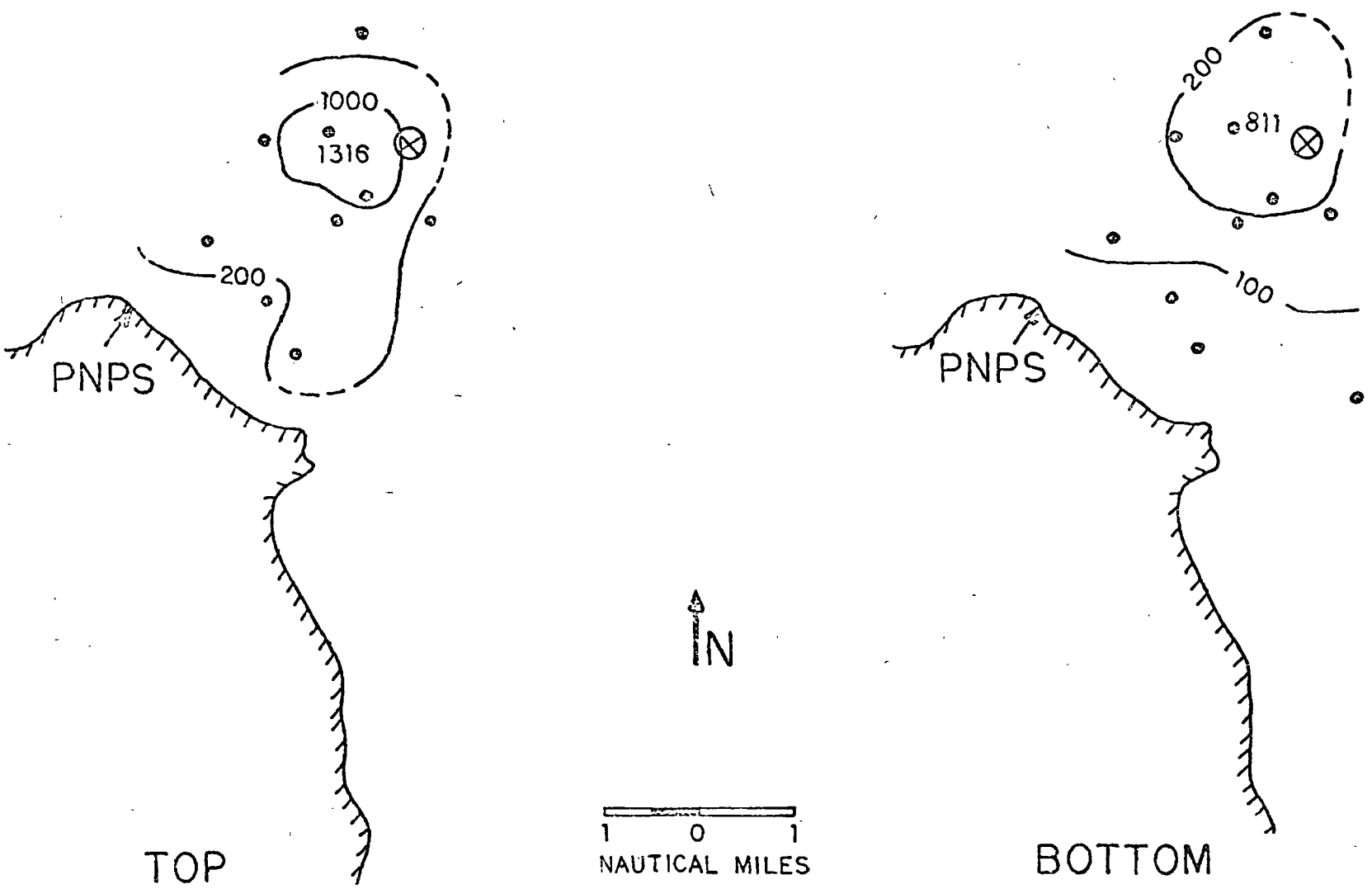


Figure 8a Experimental Results at Day D+1

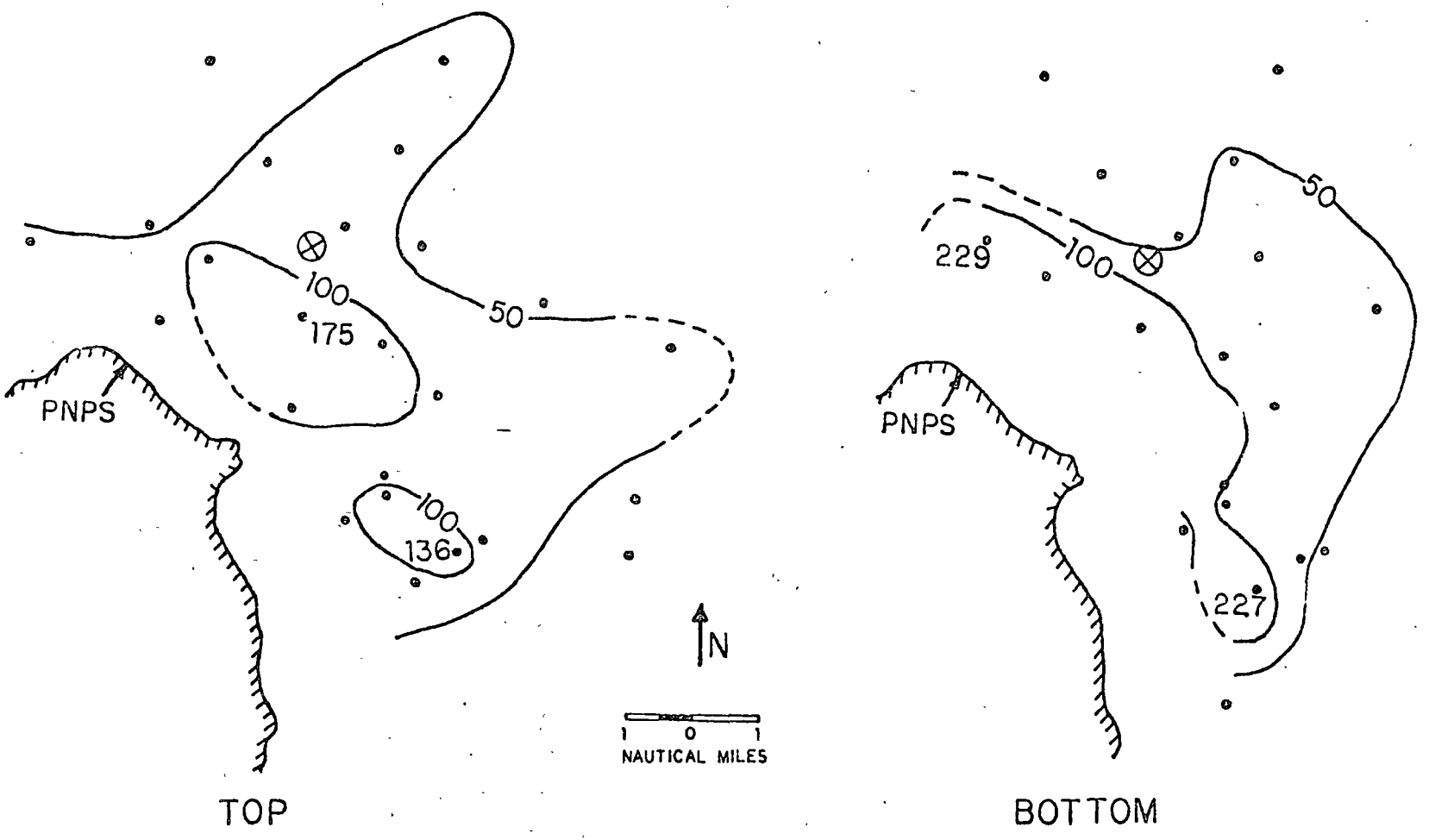


Figure 8b Experimental Results at Day D+2

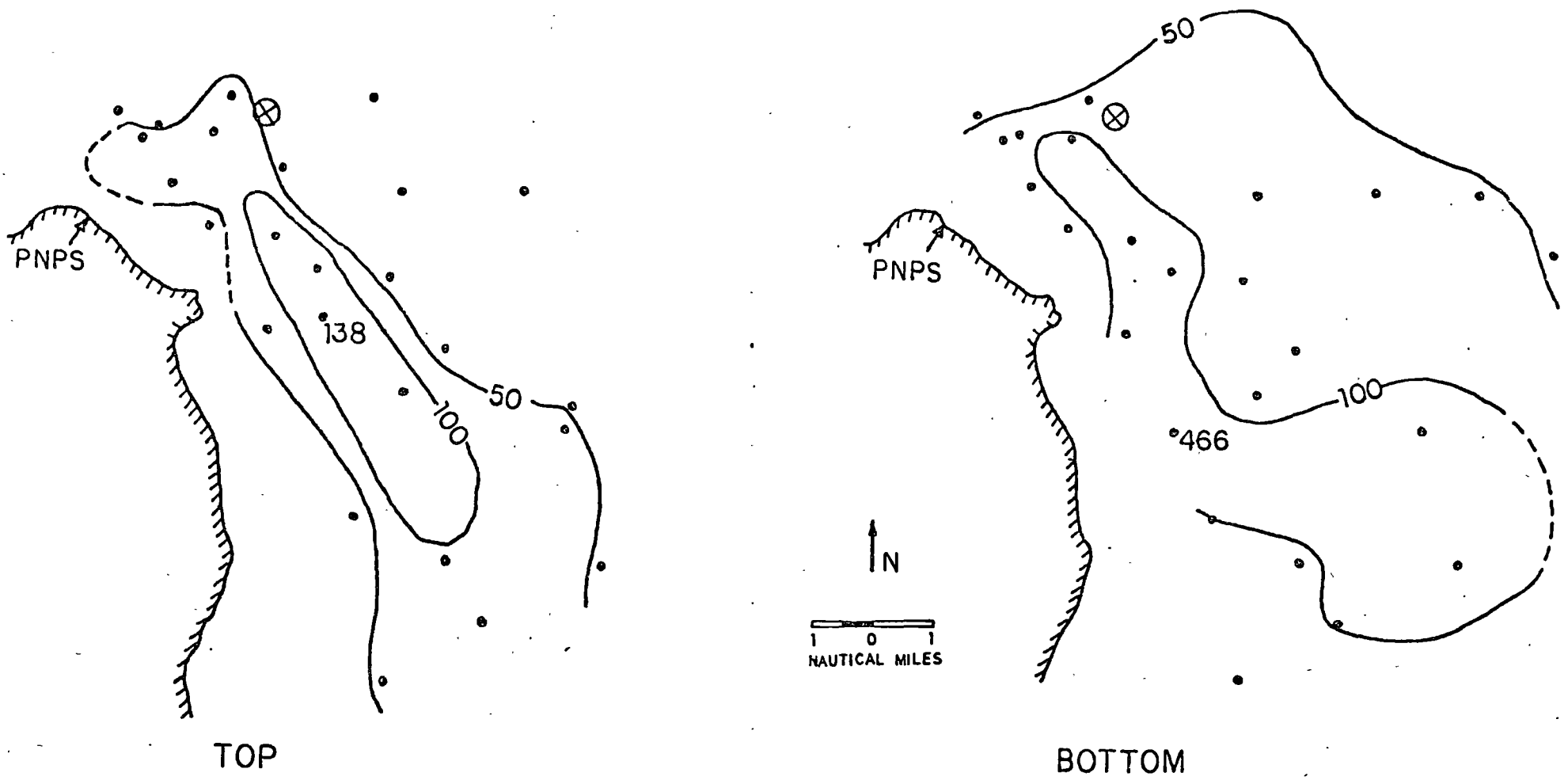


Figure 8c Experimental Results at Day D+3

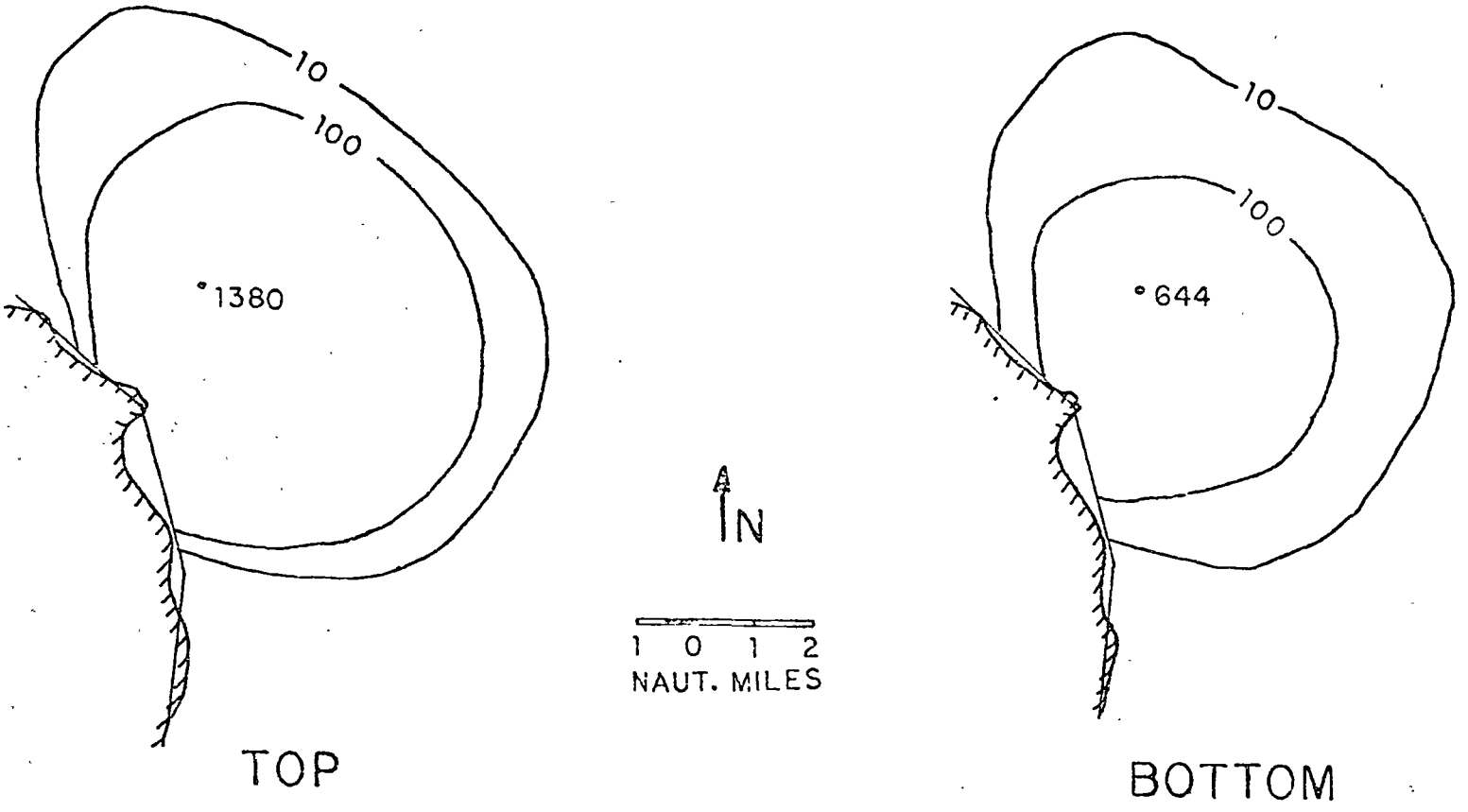


Figure 9a Computed Concentrations at Day D+1

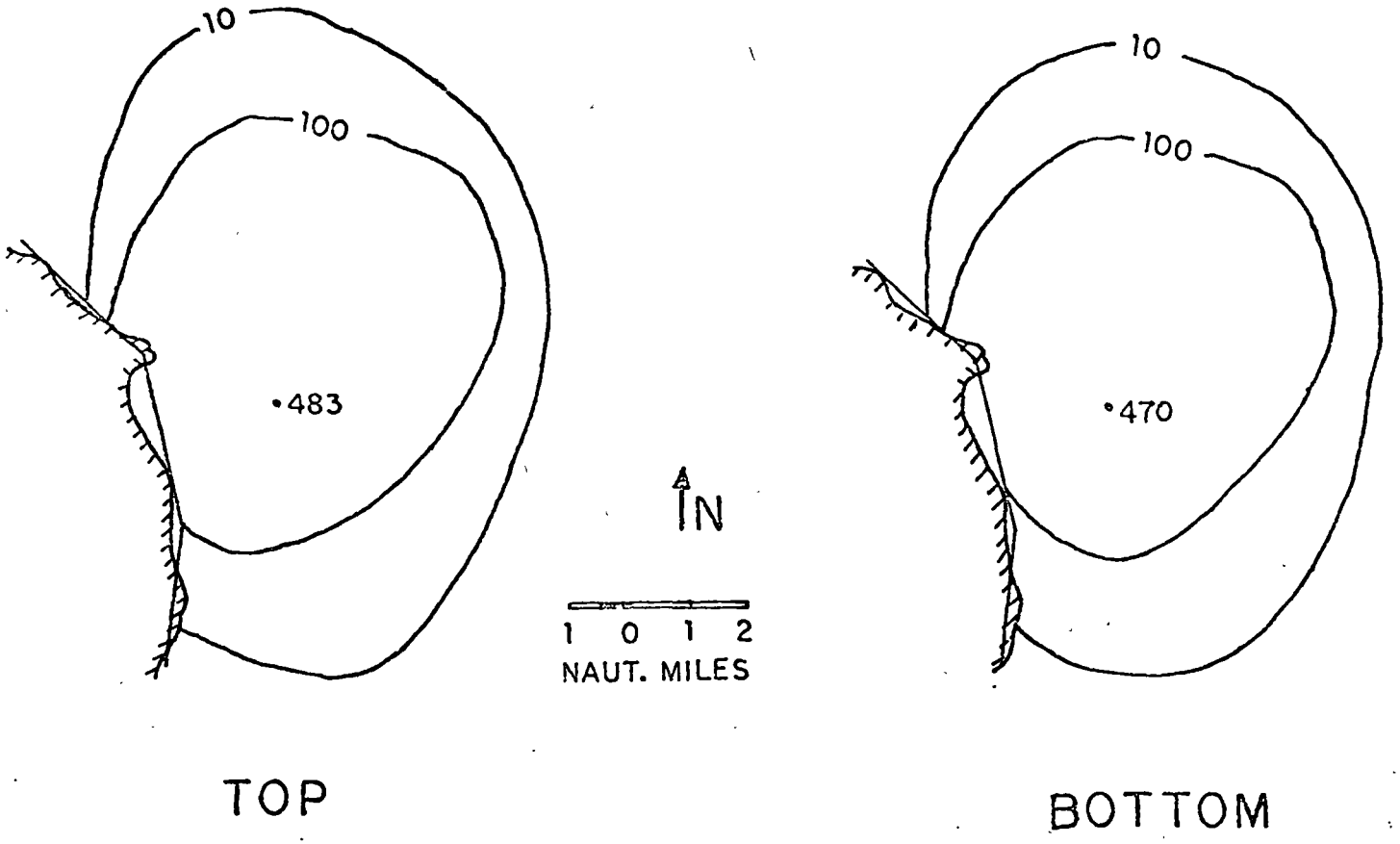


Figure 9b Computed Concentrations at Day D+2

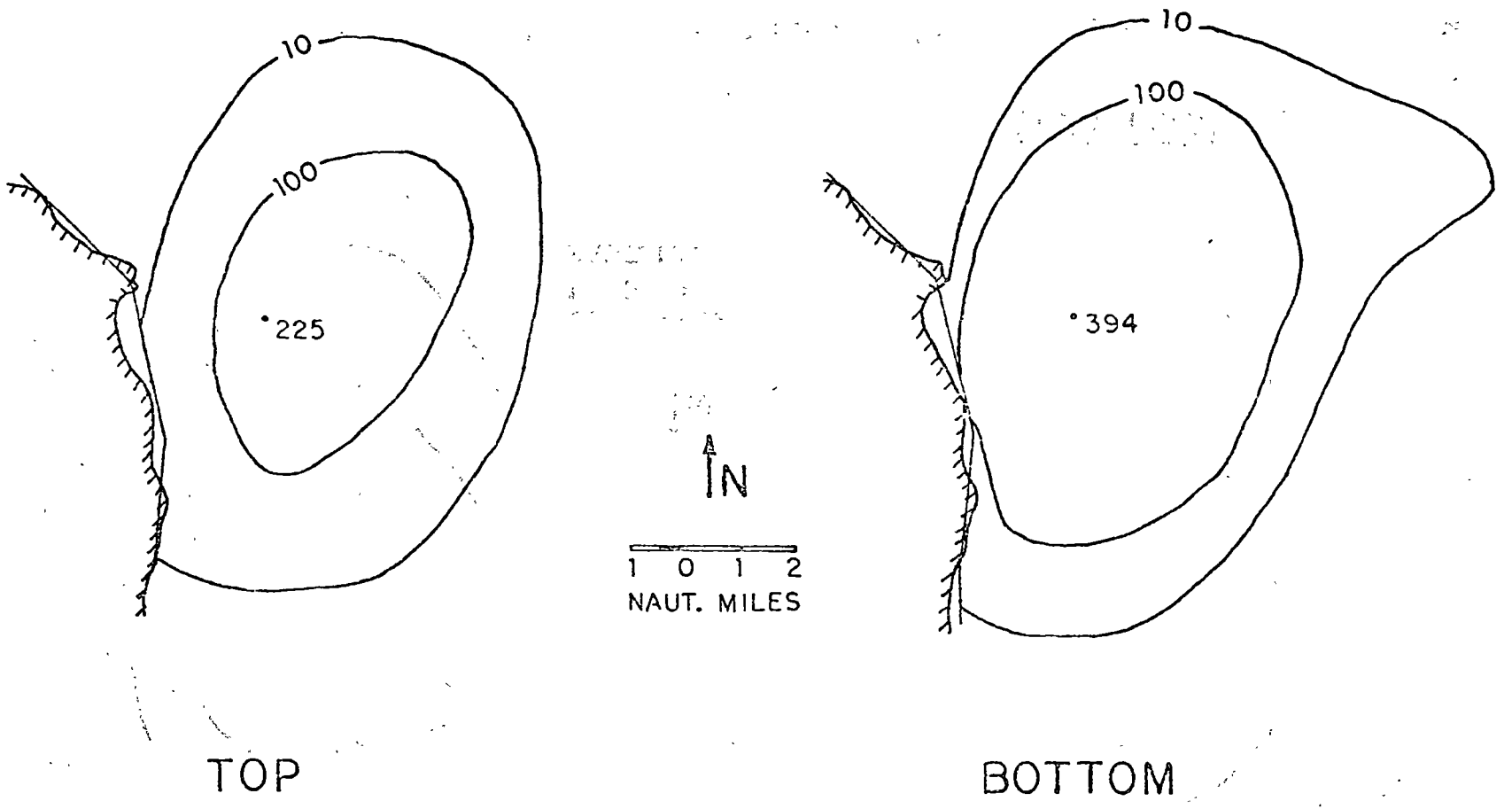
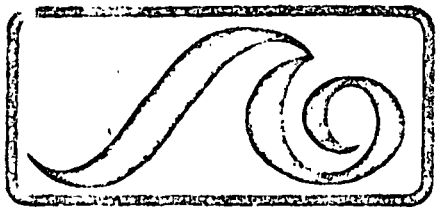


Figure 9c Computed Concentrations at Day D+3



MIT
SEA
GRANT
PROGRAM

APPLICATION OF ESTIMATION THEORY TO DESIGN OF SAMPLING PROGRAMS FOR VERIFICATION OF COASTAL DISPERSION PREDICTIONS

by
Richard De Guida
Jerome J. Connor
Bryan Pearce

Massachusetts Institute of Technology

Cambridge, Massachusetts 02139

Report No. MITSG 76-16
November 20, 1976

APPLICATION OF ESTIMATION THEORY TO DESIGN OF SAMPLING PROGRAMS FOR VERIFICATION OF COASTAL DISPERSION PREDICTIONS

by

Richard De Guida
Jerome J. Connor
Bryan Pearce

Massachusetts Institute of Technology
Sea Grant Program

Report No. MITSG 76-16
Index No. 76-316-Ccb

Acknowledgements

This report was presented at the International Conference on Finite Elements in Water Resources, Princeton University, July 1976. The research was carried on as part of the M.I.T. Sea Grant Program with support from the Office of Sea Grant in the National Oceanic and Atmospheric Administration, U.S. Department of Commerce, through grant number 04-5-158-1, and from the Massachusetts Institute of Technology.

Related Reports

Parker, Bruce B., and Bryan R. Pearce. THE RESPONSE OF MASSACHUSETTS BAY TO WIND STRESS. MITSG 75-2. Cambridge: Massachusetts Institute of Technology. February 1975.

Lassiter, Joseph B., III, James E. Soden, and Robert J. Powers. AN ASSAY OF THE MARINE RESOURCES OF MASSACHUSETTS BAY. MITSG 74-26, NTIS COM-74-11589/AS. Cambridge: Massachusetts Institute of Technology, June 1974. 58pp. $2.50.

Christodoulou, Georgios C., William F. Leimkuhler, and Arthur T. Ippen. MATHEMATICAL MODELS OF THE MASSACHUSETTS BAY; PART III: A MATHEMATICAL MODEL FOR THE DISPERSION OF SUSPENDED SEDIMENTS IN COASTAL WATERS. MITSG 74-14, NTIS COM-74-10977/AS. Cambridge: Massachusetts Institute of Technology, January 1974. 143 pp. $4.00.

Note: The preceding publications may be ordered from the National Technical Information Service, U.S. Department of Commerce, Springfield, Virginia 22151. Use the NTIS number when ordering; prices are variable.

APPLICATION OF ESTIMATION THEORY TO DESIGN OF SAMPLING PROGRAMS FOR VERIFICATION OF COASTAL DISPERSION PREDICTIONS

Richard N. De Guida

Graduate Research Assistant, Department of Civil Engineering, M.I.T., Cambridge, Mass.

Jerome J. Connor

Professor, Department of Civil Engineering, M.I.T., Cambridge, Mass.

Bryan R. Pearce

Assistant Professor, Department of Civil Engineering, Cambridge, Mass.

INTRODUCTION

Experience and engineering judgment form the basic foundation for designing sampling programs. Collection of accurate field data is required for verification of constituent (pollutant) dispersion predictions. However, the complexity of the dispersal phenomenon precludes the design of optimal sampling strategies based upon only qualitative analyses; more substantial quantitative analyses are required.

The most informative sampling strategy would require the collection of samples covering the entire spatial and temporal domains of the particular problem at sufficiently small spatial and temporal intervals to ensure the identification of all important information. A greatly reduced number of samples is usually collected due to the imposition of cost constraints. In such cases, decisions must be reached as to which samples are, and which are not, to be collected. The importance of such decisions is magnified in short-term field sampling programs. In long-term monitoring programs, one has the capabil-

ity of altering the initial strategy to improve its effectiveness as data becomes available. Due to the relatively short duration of typical field sampling programs (e.g., tracer experiments), they must be designed before the start of the sampling effort, since the results of sampling are not usually available until after the commencement of the field program. Thus it is desirable that a methodology be made available to assist in the designing of effective sampling programs.

Only within the last few years have quantitative methodologies for determining spatial and temporal sampling intervals begun to appear in the technical literature. The particular methodology of interest is based upon the concepts of Estimation Theory (specifically, Kalman-Bucy filtering). Estimation Theory refers to a variety of statistical techniques developed for determining best approximations of unknown quantities from observations (data) which are recognized as being imperfect, i.e., containing uncertainty. Kalman-Bucy filtering is a technique available for the estimation of the states of a system by the sequential extraction of information from data, as the data becomes available. It has been employed successfully in the field of navigation and guidance of spacecraft since the mid 1960's, and several investigators have recently attempted to apply these concepts to environmental pollution problems. Moore [1973] applied filtering techniques to determine the minimum monitoring frequency of certain water quality constituents for a simulated river system. Brewer and Moore [1974] extended the work of Moore [1973] to include the problem of determining the water quality constituent to be sampled and their spatial locations. Although Desalu [1974] did not directly address the monitoring design problem, he illustrated the applicability of Estimation Theory to such air pollution problems as: i) estimation of the three-dimensional distribution of pollutant concentrations from observed data, ii) identification of the diffusion coefficient and other model parameters and iii) identification of the major sources of air pollution. Pimentel [1975] illustrated that a simplified formulation results when measurements are made infrequently. This approach required ignoring the advection of the constituent; only diffusion is considered, an assumption unsuitable for estuarine areas. In addition, the important question of what is the maximum rate of sampling that can be considered as infrequent was not addressed.

A common deficiency of the above studies is the lack of effort directed at quantifying the modeling uncertainty. Although filtering concepts are straightforward, difficulty arises in their application. A major difficulty is the quantification of the modeling uncertainty. Lettenmaier [1975] considers uncertainties in tributaries, waste sources and certain parameters in his approach to design of river monitoring programs

for detection of water quality trends. The use of a steady-state one-dimensional model and temporally constant uncertainty statistics severely restricts its usefulness. The work of Dandy [1976] appears to be the most complete study published to date. He considers the design of riverine monitoring programs using a one-dimensional transient model of the advection of water quality constituents. Modeling uncertainty due to randomly varying streamflow, tributary discharges, and waste sources is considered. However, he neglects constituent dispersion and model parameter uncertainties, and uses a simplified representation of the hydrodynamics.

In this paper, the analytical framework for applying Kalman-Bucy filtering to dispersion in a coastal water body is developed. Particular emphasis is placed on quantification of the model uncertainty due to model parameters, source loadings, and velocity fields. The formulation is discretized with the Finite Element Method, and a number of comparison studies are presented.

In what follows, we outline first the filtering strategy, then describe briefly the Finite Element implementation, and lastly discuss some examples.

FILTERING CONCEPTS

Consider a linear, discrete mathematical model of the following form:

$$\underline{X}_{(t+\Delta t)} = \underline{\underline{\Phi}}_{(t)}\underline{X}_{(t)} + \underline{\underline{\beta}}_{(t)}\underline{\Psi}_{(t)} + \underline{\Omega}_{(t)} \qquad (1)$$

where

- $\underline{X}$ is a n-dimensional system state vector
- $\underline{\underline{\Phi}}$ is a n x n dimensional state transition matrix
- $\underline{\Psi}$ is a n-dimensional vector of known deterministic inputs
- $\underline{\underline{\beta}}$ is a n x n dimensional factor matrix of the deterministic input vector
- $\underline{\Omega}$ is a n-dimensional vector of model uncertainty having zero mean and covariance $Q_{M_{(t)}}$, as designated by $(0, Q_{M_{(t)}})$
- $(\ )_{t+\Delta t}$ represents the array evaluated at time $t+\Delta t$
- $(\ )_t$ represents the array evaluated at time t
- Δt is the time increment

The state-space model form of Equation 1 allows the calculation of the system state vector at time $t + \Delta t$ from the system state vector at time t. Since $\underline{\underline{\Phi}}_{(t)}$ is independent of $\underline{X}_{(t)}$, the model is linear in the independent variable $\underline{X}_{(t)}$. It is also discrete, as opposed to continuous, since it allows the computation of the dependent system state vector at only discrete times (temporally

spaced Δt units of time apart). The deterministic model would not normally include the last term, $\underline{\Omega}_{(t)}$. It is included here to signify the uncertainty in the results predicted by the deterministic model. Specification of a zero mean model uncertainty defines an unbiased model. If the model is biased, and the value of the bias is known, the model uncertainty can be represented by a deterministic bias and a zero mean random contribution.

Consider next, the following linear, discrete form of the observations:

$$\underline{Z}_{(t)} = \underline{H}_{(t)}\underline{X}_{(t)} + \underline{Y}_{(t)} \tag{2}$$

where

$\underline{Z}_{(t)}$ is a m-dimensional vector of field observations

$\underline{H}_{(t)}$ is a m x n dimensional observation matrix

$\underline{Y}_{(t)}$ is a m-dimensional vector of observation uncertainty having zero mean and covariance $Q_{o_{(t)}}$, as designated by $(0, Q_{o_{(t)}})$

The observation matrix, $\underline{H}_{(t)}$, designates the locations at which the data is collected. At each time step, a new observation matrix may be formulated, with the number of rows corresponding to the number of observations at that time. For each row, zeros (0's) appear in all columns except the column corresponding to a node of sampling; in this column, a one (1) is placed. For example, if only node 2 is measured in a 4 node system, the observation array will be:

$$[0 \quad 1 \quad 0 \quad 0]$$

Information from the model and observations can be combined by Kalman-Bucy filtering, as presented by Gelb [1974], Jazwinski [1970], and Schweppe [1973]. The first stage of the filter (i.e., prediction state) entails the extrapolation of both the state estimates and the system error covariances forward in time to the next discrete time point using the system model of Equation 1. Assuming that the model uncertainty is uncorrelated in time, the predicted system error covariance is:

$$\underline{\Sigma}_{(t+\Delta t|t)} = \underline{\Phi}_{(t)}\underline{\Sigma}_{(t|t)}\underline{\Phi}_{(t)}^{T} + \underline{Q}_{M_{(t)}} \tag{3}$$

where

$\underline{\Sigma}_{(t+\Delta t|t)}$ is a n x n dimensional predicted system error covariance matrix evaluated at time $t+\Delta t$, given measurements only up to and including time t.

$\underline{\Sigma}_{(t|t)}$ is a n x n dimensional updated system error covariance matrix evaluated at time t, given measurements up to and including time t

The above expression emphasizes that the system uncertainties are propagated through the model in a way analogous to the system states themselves. The model error covariance, $Q_{M_{(t)}}$, arises due to the error introduced in the propagation of the system errors from one time step to the next by use of the model state transition matrix.

The updated system states are obtained from the predicted system states and a linear weighting of the difference between the predicted system values and the observations as:

$$\hat{\underline{X}}_{(t+\Delta t)} = \underline{X}_{(t+\Delta t)} + \underline{\underline{K}}_{(t+\Delta t)} [\underline{Z}_{(t+\Delta t)} - \underline{\underline{H}}_{(t+\Delta t)} \underline{X}_{(t+\Delta t)}] \quad (4)$$

where

$\underline{X}_{(t+\Delta t)}$ is the n-dimensional vector of updated system states

Since minimum variance system state estimates are desired, that weighting function is computed which minimizes the trace of the predicted error covariance matrix. This weighting function, specifically called the Kalman gain matrix, is:

$$\underline{\underline{K}}_{(t+\Delta t)} = \underline{\underline{\Sigma}}_{(t+\Delta t|t)} \underline{\underline{H}}^T_{(t+\Delta t)} [\underline{\underline{H}}_{(t+\Delta t)} \underline{\underline{\Sigma}}_{(t+\Delta t|t)} \underline{\underline{H}}^T_{(t+\Delta t)} + Q_{o_{(t+\Delta t)}}]^{-1} \quad (5)$$

where

$\underline{\underline{K}}_{(t+\Delta t)}$ is the n x m dimensional Kalman gain matrix

$( )^{-1}$ indicates the inverse of the given array

$( )^T$ indicates the transpose of the given array

It is seen from Equation 5 that the Kalman gain matrix is computed from the weighting of the uncertainties in the predicted system values and the observations. With such, the updated system uncertainty is computed from:

$$\underline{\underline{\Sigma}}_{(t+\Delta t|t+\Delta t)} = [\underline{\underline{I}} - \underline{\underline{K}}_{(t+\Delta t)} \underline{\underline{H}}_{(t+\Delta t)}] \underline{\underline{\Sigma}}_{(t+\Delta t|t)} \quad (6)$$

where

$\underline{\underline{I}}$ is an n x n dimensional identity matrix

From the above, it is seen that the updated system error covariances can only be less than or equal to the predicted system error covariances. With perfect data, the system error covariances are reduced to zero at the locations of sampling. With uninformative data, the updated system error covariances will correspond exactly to the predicted system error covariances. An extremely important characteristic of the system error covariance update is its independence of the actual data values; only the statistics of the data uncertainty are required. This property allows the system error covariances to be computed before the data is made available, and thus, can be made to assist in the design of data collection programs.

To summarize the filtering process and computational requirements, the filter equations are presented in the flowchart of Figure 1. Whether data is available or not, the predicted system states and system error covariances must be calculated at each time step; the major computational cost of the filter is incurred here. In actuality, the computational difficulty and cost of filtering depends on whether the errors are white or colored (temporally invariant or correlated), and on whether the system is linear or nonlinear. The filtering algorithm presented here has made use of simplifying assumptions appropriate for linear system dynamics and temporally uncorrelated errors. For more detailed descriptions of filtering, the reader is referred to the works of Gelb [1974], Jazwinski [1970], and Schweppe [1973].

DETERMINISTIC DISPERSION MODEL

The deterministic model employed here is a vertically averaged two-dimensional finite element discretization which is applicable when the velocity and concentration vary slowly over the water column, i.e., for well mixed conditions. We have restricted the treatment to a vertically averaged formulation since our objective was to investigate the computational feasibility of applying filtering techniques and a three-dimensional treatment would be premature at this time.

Integrating the general convective diffusion equation over the water column results in the following governing equation (see Leimkuhler [1974] for details):

$$\frac{\partial}{\partial t} C + \frac{\partial}{\partial x} (\bar{u}\, C) + \frac{\partial}{\partial y} (\bar{v}\, C) = - \frac{\partial}{\partial x} Q_x - \frac{\partial}{\partial y} Q_y + \tilde{S}_i + \tilde{S}_s + \tilde{S}_b \qquad (7)$$

where

C is the depth integrated concentration,

$$C = \rho \bar{c}\, h$$

ρ is the mass density of the constituent and water mixture

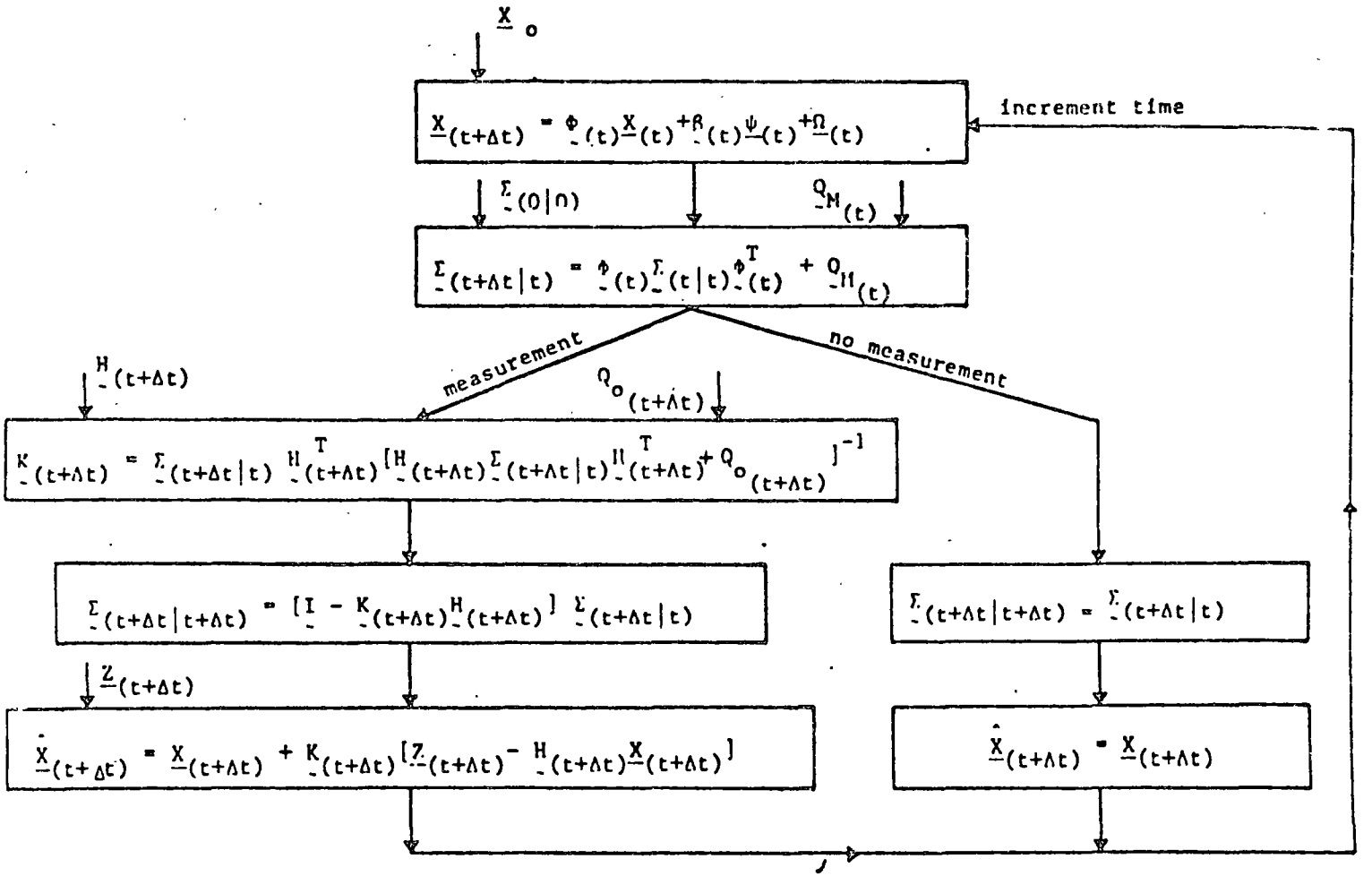


Figure 1 Schematic Diagram of Kalman Filtering Process

(from Pearce, et al. [1975])

$\bar{c}$, $\bar{u}$, $\bar{v}$ are the depth-averaged values of concentration and horizontal velocity components

h is the total height of water column

$\tilde{S}_i$ is the constituent mass input rate per unit projected area

$\tilde{S}_s$, $\tilde{S}_b$ are the normal source loading flux components through the surface and bottom of the water column

The flux terms are approximated with isotropic Fickian dispersion expressions,

$$Q_x = - \rho h E_{x,y} \frac{\partial \bar{c}}{\partial x} \tag{8}$$

$$Q_y = - \rho h E_{x,y} \frac{\partial \bar{c}}{\partial y} \tag{9}$$

where $E_{x,y}$ is the isotropic dispersion coefficient

For the particular case of settling of discrete particles (e.g., sphalerite tracer particles, suspended sediment, etc.), the source and sink terms are simplified by assuming a first order decay rate due to the settling velocity of the particles, and constant concentration through the layer. This yields

$$\tilde{S}_i + \tilde{S}_s + \tilde{S}_b = \tilde{S}_i - \rho w_s \bar{c} \tag{10}$$

where w_s is the particle settling velocity.

For coastal problems, the concentration is specified as C = 0 on the ocean boundary when the boundary is far from the plume. The normal dispersive flux is specified as $Q_n = 0$ along the land boundary. If the plume intersects the ocean boundary, the normal dispersive flux can be prescribed as being equal to zero provided that the concentration is constant outside the domain.

Equation (7) is transformed to its symmetrical weak form and the Finite Element spatial discretization is applied. The details are presented by Leimkuhler, et al (1975), and we list here only the final form of the governing equation:

$$\underset{\sim}{M} \frac{\partial}{\partial t} \underline{C} + \underset{\sim}{A} \cdot \underline{C} + E_{x,y} \cdot \underset{\sim}{K} \cdot \underline{C} + w_s \cdot \underset{\sim}{D} \cdot \underline{C} - \underline{S} + \underline{F} = 0 \tag{11}$$

where A contains the advective terms, K defines the dispersion component, D refers to decay, S contains the source loading, and F represents the dispersive boundary flux term.

The trapezium method is employed to propagate the solution in time. In the deterministic case, the scheme is relatively inexpensive since the state transition matrix, Φ, does not have to be generated. However, it is required for the covariance propagation If advection is treated with the "pure" trapezium, the generation of Φ would require matrix factorization at each time step. To reduce this effort, an Eulerian approximation for advection is introduced and the solution is propogated with

$$[\underset{\sim}{M} + \frac{\Delta t}{2}(E_{x,y} \cdot \underset{\sim}{K} + w_s \cdot \underset{\sim}{D})]\underline{C}_{n+1} = [\underset{\sim}{M} - \frac{\Delta t}{2}(E_{x,y} \cdot \underset{\sim}{K} + w_s \cdot \underset{\sim}{D}) - \Delta t \quad \underset{\sim}{A}_n]\underline{C}_n + \frac{\Delta t}{2}(\underline{\hat{S}}_{n+1} + \underline{\hat{S}}_n) \tag{12}$$

where

$(\ )_n$ designates the given array evaluated at the discrete time point, t_n

Δt is the time increment

The Euler approximation for advection decreases the stability limit but this is usually not a problem for coastal dispersion.

QUANTIFICATION OF DISPERSION MODELING UNCERTAINTY

The major effort required in applying Kalman-Bucy filtering to coastal dispersion problems is quantification of the modeling uncertainty. Since coastal dispersion generally involves a large number of unknowns, only a first-order uncertainty analysis is feasible. In first-order analyses, each variable is considered to be a random function in which the mean represents the best estimate of the variable, and the variance quantifies the uncertainty in the estimate.

To compute the uncertainty in the predicted concentrations, due to parameter and input uncertainties, the deterministic model is expanded in a Taylor series about the mean values of the variables. Retaining only first-order terms, results in the following equation which defines the propagation of the uncertainty in concentration:

$$\left[\underset{\sim}{M} + \frac{\Delta t}{2}\left(E_{x,y}\cdot\underset{\sim}{K} + w_s\cdot\underset{\sim}{D}\right)\right]\tilde{\underset{\sim}{C}}_{t+\Delta t}$$

$$= \left[\underset{\sim}{M} - \frac{\Delta t}{2}\left(E_{x,y}\cdot\underset{\sim}{K} + w_s\cdot\underset{\sim}{D}\right) - \Delta t\cdot\underset{\sim}{A}_t\right]\tilde{\underset{\sim}{C}}_t$$

$$- \left[\frac{\Delta t}{2}\left(\tilde{E}^{x,y}_t\cdot\underset{\sim}{K} + \tilde{w}_{s_t}\cdot\underset{\sim}{D}\right) + \Delta t\cdot\tilde{\underset{\sim}{A}}_t\right]\underset{\sim}{C}_t$$

$$- \left[\frac{\Delta t}{2}\left(\tilde{E}^{x,y}_{t+\Delta t}\cdot\underset{\sim}{K} + \tilde{w}_{s_{t+\Delta t}}\cdot\underset{\sim}{D}\right)\right]\underset{\sim}{C}_{t+\Delta t}$$

$$+ \frac{\Delta t}{2}\left[\left(\tilde{\underset{\sim}{S}}\right)_t + \left(\tilde{\underset{\sim}{S}}\right)_{t+\Delta t}\right] \qquad (13)$$

where $\tilde{(\ )}_t$ represents the uncertainty in the given variable at time t

Our representation of the model parameters and inputs is equivalent to considering the uncertainty about the mean value as a zero mean process. The isotropic dispersion coefficient and first-order decay rate uncertainties are interpreted as

$$\tilde{E}^{x,y} \sim (0, \sigma_E{}^2{}_{x,y})$$

$$\tilde{w}_s \sim (0, \sigma_{w_s}{}^2)$$

where σ_x^2 represents the variance of the uncertainty in the variable x

Representation of the model inputs uncertainty creates more difficulty. For multi-location source discharges, each discharge would normally have its own characteristic level of uncertainty. However, to simplify, the loading is expressed in terms of a single loading parameter and a vector defining the spatial distribution of the loading as,

$$\hat{\underline{S}}_t = \lambda_t \underline{R}_t$$

where

λ_t is the loading parameter

$\underline{R}_t$ is a vector describing the geographic locations of the loadings

If only one source location exists, such as in most tracer field experiments, the above expression is exact. The uncertainty in the loading parameter is represented as a random function with zero mean and prescribed variance,

$$\tilde{\lambda}_t \sim (0, \sigma_{\lambda_t}^2)$$

where $\tilde{\lambda}_t$ is the source loading uncertainty

$\sigma_{\lambda_t}^2$ is the variance of the source loading uncertainty

The flexibility of handling temporally and spatially variant velocity fields creates difficulty in representation of the advection field uncertainty. Our approach is based on formulating the uncertainty at the element level similar to the formulation of the element advection matrix of the deterministic dispersion model.

Equation 13 shows that the effect of velocity uncertainty on the uncertainty of the predicted model concentrations is determined from

$$[\underset{\sim}{M} + \frac{\Delta t}{2} (E_{x,y} \cdot \underset{\sim}{K} + w_s \cdot \underset{\sim}{D})] \tilde{\underline{C}}_{t+\Delta t} = -\Delta t \cdot \tilde{\underset{\sim}{A}}_t \cdot \underline{C}_t \qquad (14)$$

The advection uncertainty term is decomposed into influence matrices and vectors of the x and y component velocity field uncertainties as

$$-\Delta t \cdot \tilde{\underline{A}}_t \cdot \underline{C}_t = \underline{A}_t^{(u)} \cdot \tilde{\underline{u}}_t + \underline{A}_t^{(v)} \cdot \tilde{\underline{v}}_t \tag{15}$$

where $\tilde{\underline{u}}_t$ is the vector of x component velocity field uncertainties

$\tilde{\underline{v}}_t$ is the vector of y component velocity field uncertainties

Using Equations 14 and 15, the following factor matrices can be defined (details are presented in DeGuida [1976]):

$$\underline{\phi}_{u,t} = [\underline{M} + \frac{\Delta t}{2} (E_{x,y} \cdot \underline{K} + w_s \cdot \underline{D})]^{-1} \cdot \underline{A}_t^{(u)} \tag{16}$$

$$\underline{\phi}_{v,t} = [\underline{M} + \frac{\Delta t}{2} (E_{x,y} \cdot \underline{K} + w_s \cdot \underline{D})]^{-1} \cdot \underline{A}_t^{(v)} \tag{17}$$

Collecting the various uncertainty contributions, the two-dimensional model uncertainty expands to:

$$\begin{aligned} \tilde{\underline{C}}_{t+\Delta t} = {} & \underline{\phi}_{c,t+\Delta t} \tilde{\underline{C}}_t \\ & + \tilde{E}_t^{x,y} \cdot \underline{\phi}_{E_{x,y},t} + \tilde{E}_{t+\Delta t}^{x,y} \cdot \underline{\phi}_{E_{x,y},t+\Delta t} \\ & + \tilde{w}_{s_t} \cdot \underline{\phi}_{w_s,t} + \tilde{w}_{s_{t+\Delta t}} \cdot \underline{\phi}_{w_s,t+\Delta t} \\ & + \tilde{\lambda}_t \cdot \underline{\phi}_{\lambda,t} + \tilde{\lambda}_{t+\Delta t} \cdot \underline{\phi}_{\lambda,t+\Delta t} \\ & + \underline{\phi}_{u,t} \cdot \tilde{\underline{u}}_t + \underline{\phi}_{v,t} \cdot \tilde{\underline{v}}_t \end{aligned} \tag{18}$$

where

$$\underline{\phi}_{c,t+\Delta t} = [\underline{M} + \frac{\Delta t}{2} (E_{x,y} \cdot \underline{K} + w_s \cdot \underline{D})]^{-1}$$

$$[\underline{M} - \frac{\Delta t}{2} (E_{x,y} \cdot \underline{K} + w_s \cdot \underline{D}) - \Delta t \cdot \underline{A}_t]$$

$$\underset{\sim}{\phi}_{E_{x,y},t} = [\underset{\sim}{M} + \frac{\Delta t}{2}(E_{x,y} \cdot \underset{\sim}{K} + w_s \cdot \underset{\sim}{D})]^{-1} \; (-\frac{\Delta t}{2} \cdot \underset{\sim}{K} \cdot \underset{\sim}{C}_t)$$

$$\underset{\sim}{\phi}_{E_{x,y},t+\Delta t} = [\underset{\sim}{M} + \frac{\Delta t}{2}(E_{x,y} \cdot \underset{\sim}{K} + w_s \cdot \underset{\sim}{D})]^{-1}$$

$$(-\frac{\Delta t}{2} \cdot \underset{\sim}{K} \cdot \underset{\sim}{C}_{t+\Delta t})$$

$$\underset{\sim}{\phi}_{w_s,t} = [\underset{\sim}{M} + \frac{\Delta t}{2}(E_{x,y} \cdot \underset{\sim}{K} + w_s \cdot \underset{\sim}{D})]^{-1} \; (-\frac{\Delta t}{2} \cdot \underset{\sim}{D} \cdot \underset{\sim}{C}_t)$$

$$\underset{\sim}{\phi}_{w_s,t+\Delta t} = [\underset{\sim}{M} + \frac{\Delta t}{2}(E_{x,y} \cdot \underset{\sim}{K} + w_s \cdot \underset{\sim}{D})]^{-1}$$

$$(-\frac{\Delta t}{2} \cdot \underset{\sim}{D} \cdot \underset{\sim}{C}_{t+\Delta t})$$

$$\underset{\sim}{\phi}_{\lambda,t} = [\underset{\sim}{M} + \frac{\Delta t}{2}(E_{x,y} \cdot \underset{\sim}{K} + w_s \cdot \underset{\sim}{D})]^{-1} \; (\frac{\Delta t}{2} \cdot \underset{\sim}{R}_t)$$

$$\underset{\sim}{\phi}_{\lambda,t+\Delta t} = [\underset{\sim}{M} + \frac{\Delta t}{2}(E_{x,y} \cdot \underset{\sim}{K} + w_s \cdot \underset{\sim}{D})]^{-1} \; (\frac{\Delta t}{2} \cdot \underset{\sim}{R}_{t+\Delta t})$$

$\underset{\sim}{\phi}_{u,t}$ and $\underset{\sim}{\phi}_{v,t}$ are as defined in Equations 16 and 17.

The propagation of the variance is obtained by squaring the uncertainty and taking the expected value. The two-dimensional form is, assuming stationary random processes (i.e., time invariant statistics of the uncertainty) (see DeGuida [1976]):

$$\underset{\sim}{\Gamma}_{c,t+\Delta t} = \underset{\sim}{\phi}_{c,t+\Delta t} \underset{\sim}{\Gamma}_{c,t} \underset{\sim}{\phi}^T_{c,t+\Delta t}$$

$$+ \sigma^2_{E_{x,y}} [\underset{\sim}{\phi}_{E_{x,y},t} \underset{\sim}{\phi}^T_{E_{x,y},t} + \underset{\sim}{\phi}_{E_{x,y},t+\Delta t} \underset{\sim}{\phi}^T_{E_{x,y},t+\Delta t}$$

$$+ (\underset{\sim}{\phi}_{c,t+\Delta t} \underset{\sim}{\phi}_{E_{x,y},t} \underset{\sim}{\phi}^T_{E_{x,y},t}) + (\underset{\sim}{\phi}_{c,t+\Delta t} \underset{\sim}{\phi}_{E_{x,y},t} \underset{\sim}{\phi}^T_{E_{x,y},t})^T]$$

$$+ \sigma^2_{w_s} [\underset{\sim}{\phi}_{w_s,t} \underset{\sim}{\phi}^T_{w_s,t} + \underset{\sim}{\phi}_{w_s,t+\Delta t} \underset{\sim}{\phi}^T_{w_s,t+\Delta t}$$

$$+ \left(\underset{\sim}{\phi}_{c,t+\Delta t}\ \underset{\sim}{\phi}_{w_s,t}\ \underset{\sim}{\phi}^T_{w_s,t}\right) + \left(\underset{\sim}{\phi}_{c,t+\Delta t}\ \underset{\sim}{\phi}_{w_s,t}\ \underset{\sim}{\phi}^T_{w_s,t}\right)^T]$$

$$+ \sigma_\lambda^2\ [\underset{\sim}{\phi}_{\lambda,t}\ \underset{\sim}{\phi}^T_{\lambda,t} + \underset{\sim}{\phi}_{\lambda,t+\Delta t}\ \underset{\sim}{\phi}^T_{\lambda,t+\Delta t}$$

$$+ \left(\underset{\sim}{\phi}_{c,t+\Delta t}\ \underset{\sim}{\phi}_{\lambda,t}\ \underset{\sim}{\phi}^T_{\lambda,t}\right) + \left(\underset{\sim}{\phi}_{c,t+\Delta t}\ \underset{\sim}{\phi}_{\lambda,t}\ \underset{\sim}{\phi}_{\lambda,t}{}^T\right)^T]$$

$$+ \underset{\sim}{\phi}_{u,t}\ \underset{\sim}{\Gamma}_u\ \underset{\sim}{\phi}^T_{u,t} + \underset{\sim}{\phi}_{v,t}\ \underset{\sim}{\Gamma}_v\ \underset{\sim}{\phi}^T_{v,t} \qquad (19)$$

where

$\underset{\sim}{\Gamma}_{c,t}$ is the variance of the model predicted concentrations at time t

$\underset{\sim}{\Gamma}_u$ is the variance of the x-directional depth averaged velocity components

$\underset{\sim}{\Gamma}_v$ is the variance of the y-directional depth averaged velocity components

SAMPLING EFFECTIVENESS DETERMINATION

Once the modeling uncertainty is quantified, a Kalman filtering algorithm is developed to quantify the effectiveness of sampling, as has been shown by previous investigators (e.g., Moore (1973), Brewer and Moore (1974), Pimentel (1975), and Dandy (1976)). With Kalman filtering, all that needs to be specified for the update of the system uncertainty are the observation matrices defining the locations of measurements in time and the statistics of the uncertainty in those measurements. If an estimate of the measurement uncertainty is available, the system uncertainty can be propagated in time, considering different observation matrices, i.e., different sampling strategies. Comparing the system uncertainty allows one to evaluate the potential effectiveness of various sampling strategies before the actual experiment is performed.

In designing sampling strategies by Kalman filtering, careful analysis is required in defining the observation matrices. The formulation as stated is entirely general, such that an infinite number of possible sampling networks (i.e., definition of the time-variant spatial locations of sampling) can be analyzed, if so desired. However, specific characteristics of each problem will normally limit the possible number

of sampling strategies to be tested. Of concern might be such factors as budgets allotted for sampling, required level of information from sampling, restricted sampling days and/or hours of sampling (e.g., only sunlight hours), political boundaries, certain legal aspects, and so on. All these factors, and many more, will probably influence the selection of possible sampling strategies. Of extreme importance here is the use of experience in sampling and engineering judgment.

For observations to be informative, the uncertainty of the observations must be less than the uncertainty in the predicted concentrations, i.e., the updated system uncertainty must be less than the predicted system uncertainty. It is therefore natural to choose the sampling strategy of minimum system uncertainty (i.e., the minimum system error covariance matrix). However, as the duration of time increases after the last observation has been made, the system uncertainty increases until finally no reduction in the system uncertainty is noticed. Therefore, at different times, different measures of the effectiveness of sampling would be obtained. To compute the sampling effectiveness over the entire time duration of the experiment, the reduction in the system error covariance matrix (i.e., predicted system error covariance matrix minus updated system error covariance matrix) is calculated at each time a sample is collected. Since the majority of the reduction occurs in the uncertain variances (i.e., diagonal elements of the system error covariance matrix), only the reductions in the trace are computed. Summation over time of these component reductions leads to a total measure of sampling effectiveness. Maximization of the total reductions of the system uncertainty is therefore an appropriate measure of sampling effectiveness for the specific problem of tracer experiments.

Some may criticize the previously described optimality criterion for the simplistic way of defining the feasible set of sampling strategies, i.e., observation matrices. In reality, even though a particular sampling strategy may not satisfy all the constraints, the penalty incurred in the constraint violation may be so small that the design will be more effective than all the others tested which satisfy the imposed constraints. These problems can be avoided by the definition of a utility function which could be made to reflect the value of sampling in light of all the complicating factors. The criterion for determining the most effective sampling network would be that network which maximizes the expected utility. Such a maximization of the expected utility has become a traditional objective in Bayesian statistical decision theory.

A major disadvantage of defining the sampling effectiveness as

maximization of the expected utility is the difficulty of expressing the utility in mathematical form. It is often very difficult to quantify certain characteristics of the problem; it may be practically feasible for only very special cases. Therefore, in light of the necessity to develop simpler evaluation criteria, evaluation of only feasible sampling strategies, as previously described, appears to be the most appropriate for the purposes of this study.

RESULTS

For purposes of illustrating the usefulness of the filtering algorithm for evaluating sampling effectiveness, some results for the one-dimensional modeling of a channel are presented first. The finite element grid used is shown in Figure 2. A constant depth of 1 meter is used. Contaminant is continuously injected at the source node. Zero concentrations are specified at the extreme ends of the grid (i.e., at $x = 0$ and $x = 3$ meters). A time increment of 0.1 seconds is used in the model. The mean values and standard deviations of the model parameters and inputs are:

Parameter	Mean Value	Standard Deviation
longitudinal dispersion coefficient	0.01 meters^2/sec	0.005 meters^2/sec
first order decay rate	0.2/sec	0.1/sec
Input		
longitudinal flow velocity	0.05 meters/sec	0.01 meters/sec
continuous source loading rate	1 gram/sec	0.1 gram/sec

In addition, the standard deviation of the measurement uncertainty is taken as 0.01 grams/meter^3.

Examination of the deterministic solution shows that the peak concentration occurs at the source location. Therefore, the first sampling strategy evaluated consisted of sampling at the source discharge location (node 9 in Figure 2) every second after the start of discharge. Since the trace of the error covariance matrix is the desired measure of sampling effectiveness, a plot of the trace of the error covariance matrix versus time is presented in Figure 3. The solid line in the

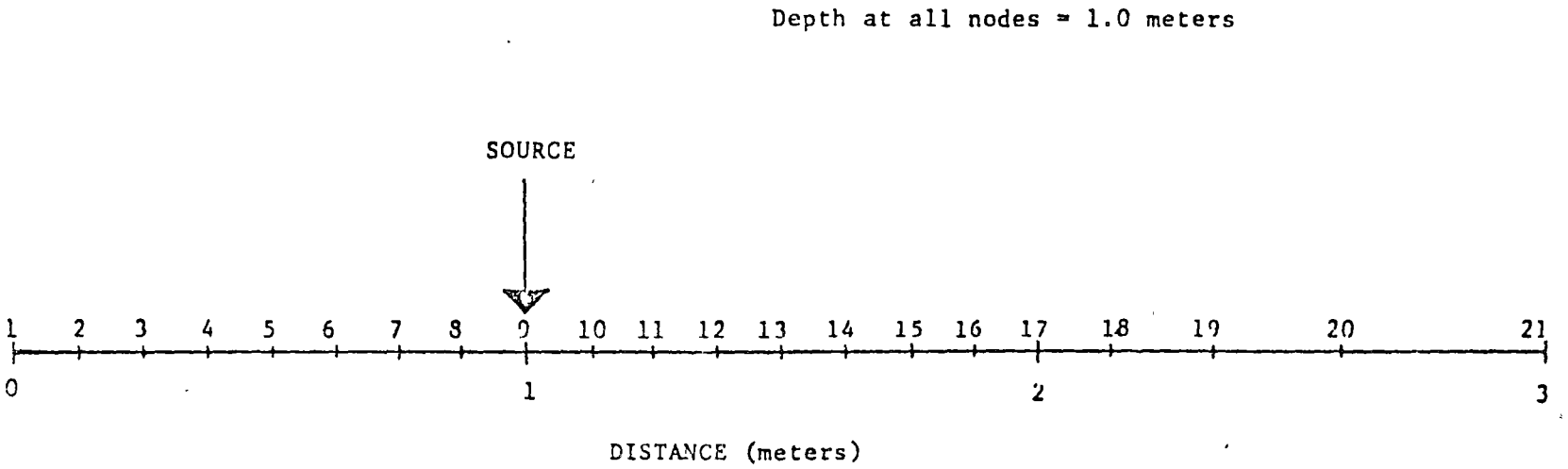


Figure 2 Finite Element Grid Used for One-dimensional Dispersion Model Simulations

figure represents the modeling uncertainty (i.e., system uncertainty with no measurements); the dashed line represents the system uncertainty as measurements are taken. Each measurement reduces the system uncertainty at the time of measurement. However, as time progresses, the system uncertainty increases, but not quite to the level corresponding to no measurement. This indicates that the measurements are informative. A measure of their value is the total reduction in system uncertainty over time.

Figure 4 shows the variation of system uncertainty with distance from the source. The solid line defines the standard deviation of the system uncertainty immediately before the measurement is taken at 10 seconds; the dashed line is the "corrected" distribution. The effect of the measurement is quite local, reducing the system uncertainty to essentially the measurement uncertainty at the observation point but diminishing rapidly away from the source.

Figure 5 illustrates the effect of sampling every second after the start of discharge at the source node and half a meter downstream (i.e., nodes 9 and 13), as measured by the trace of the error covariance matrix. Figure 6 shows the reduction of the system uncertainty as a function of distance from the source at 10 seconds after start of discharge. The additional downstream observation point reduces the covariance trace, and also increases the spatial extent of the correction. The effect of increasing the sampling frequency while sampling at only the source node is illustrated by comparing Figures 3 and 7. Figure 7 illustrates the effect of taking a sample every half second.

Considerable interest in the Massachusetts Bay environment has been expressed in connection with a once proposed offshore sand and gravel dredging project called NOMES (New England Offshore Mining Environmental Study). Such interest has motivated this study, and it seemed logical to attempt a simulation of the NOMES dispersion experiment.

The finite element grid of Massachusetts Bay is shown in Figure 8. The ability to use elements of different sizes and shapes affords the flexibility required to model such complex geometric configurations as Massachusetts Bay. The dump site of tracer particles (sphalerite) is indicated by the starred area. Depths at the nodal points are taken from the Coast and Geodetic Survey bathymetric chart 0808N-50.

Velocity time histories were generated with a two-dimensional finite element circulation model, CAFE (Circulation Analyses by Finite Elements) (see Wang and Connor [1975]) using simulated tidal input and actual wind conditions collected during

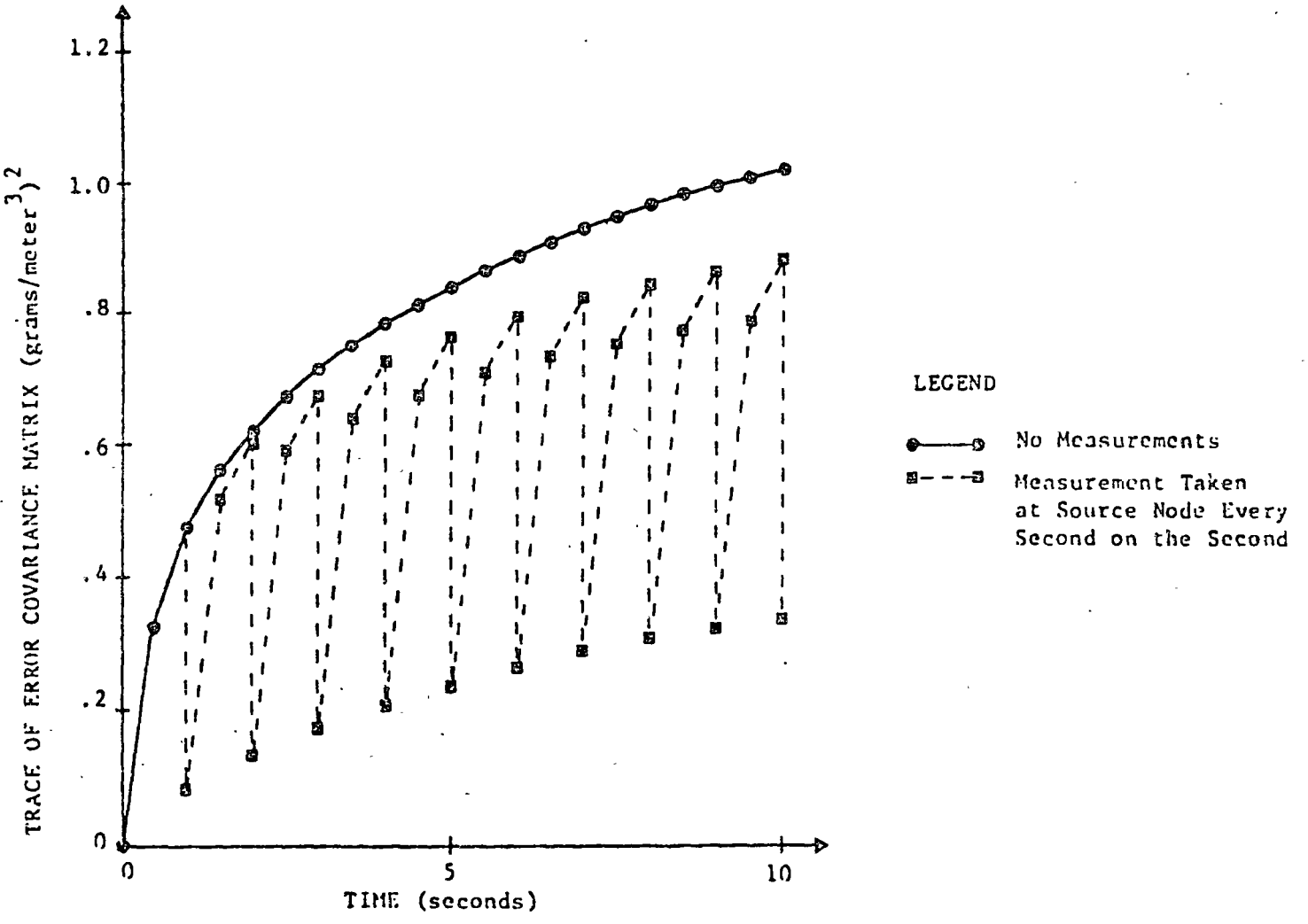


Figure 3 Temporal Effect of Sampling at Source Node Every Second

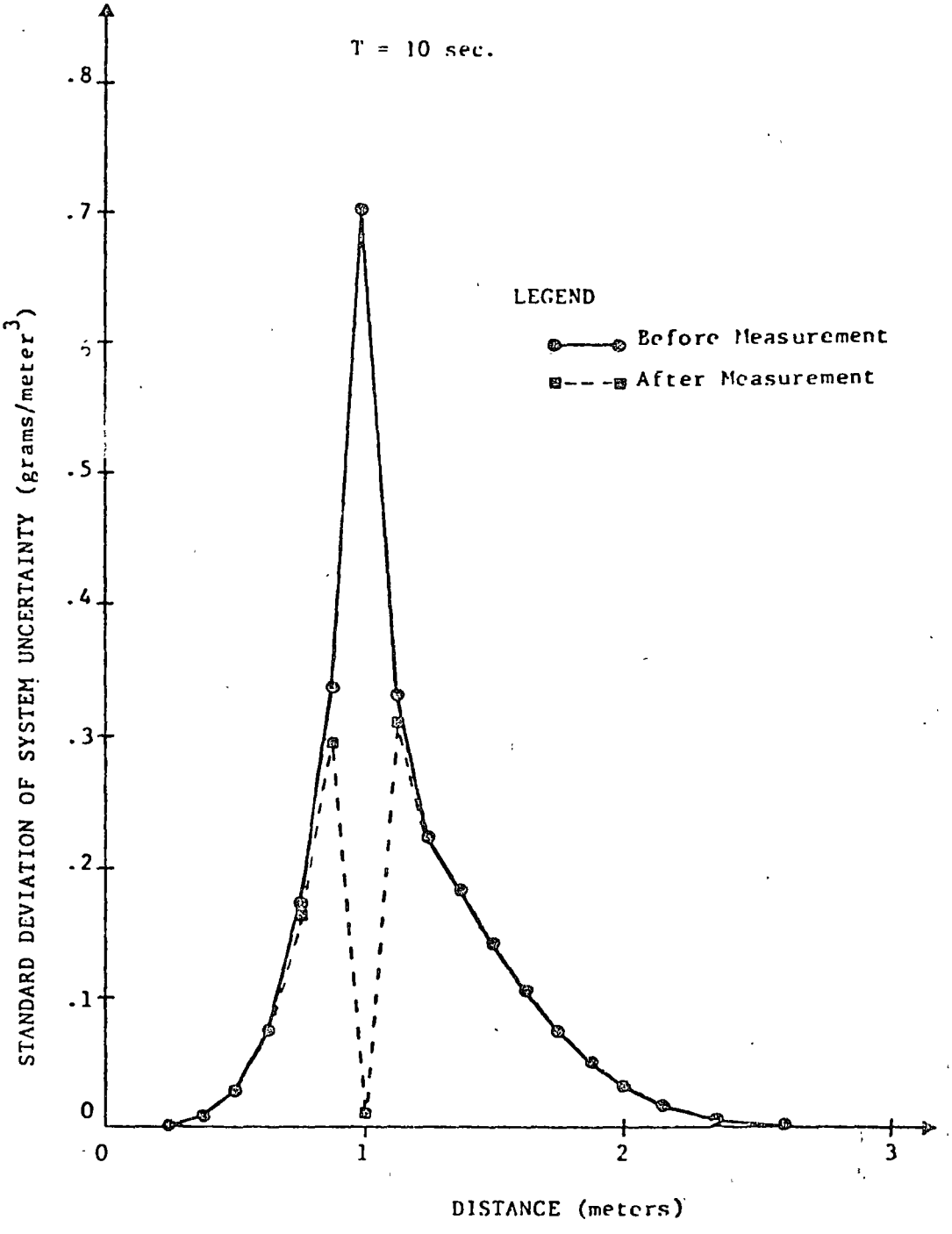


Figure 4 Spatial Effect of Sampling at Source Node for T = 10 Seconds

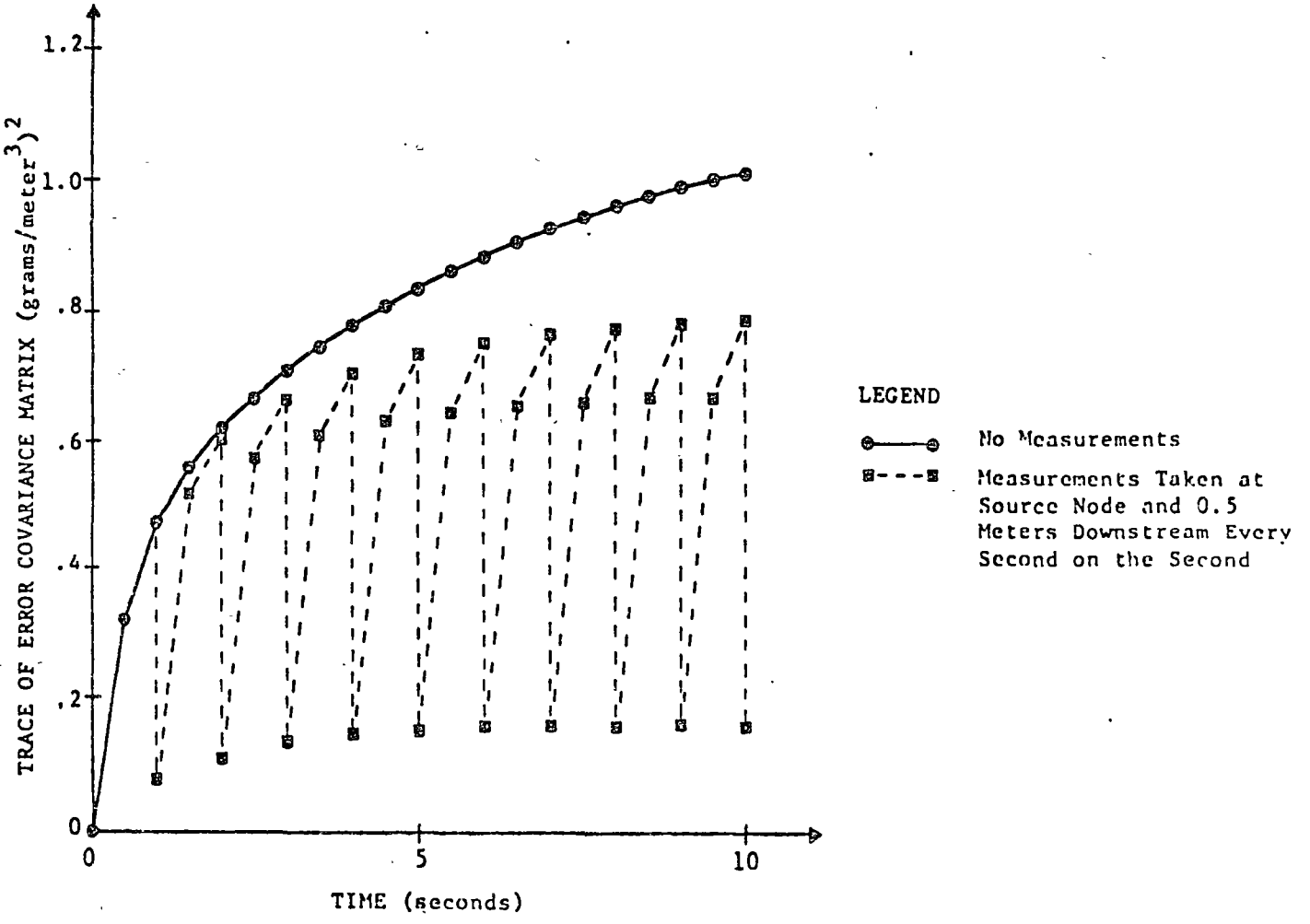


Figure 5 Temporal Effect of Sampling at Source Node and 0.5 Meters Downstream Every Second

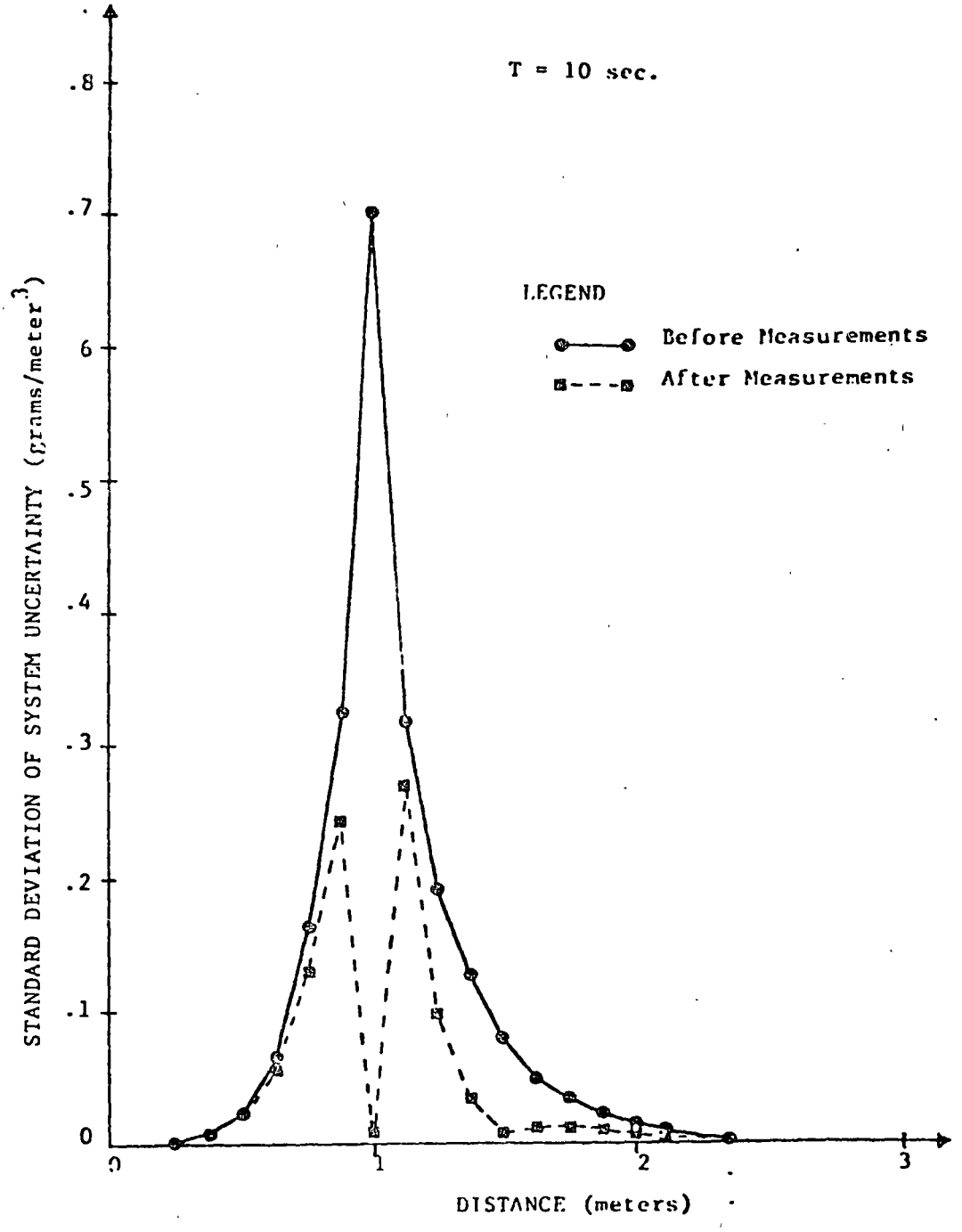


Figure 6 Spatial Effect of Sampling at Source Node and 0.5 Meters Downstream for T = 10 seconds

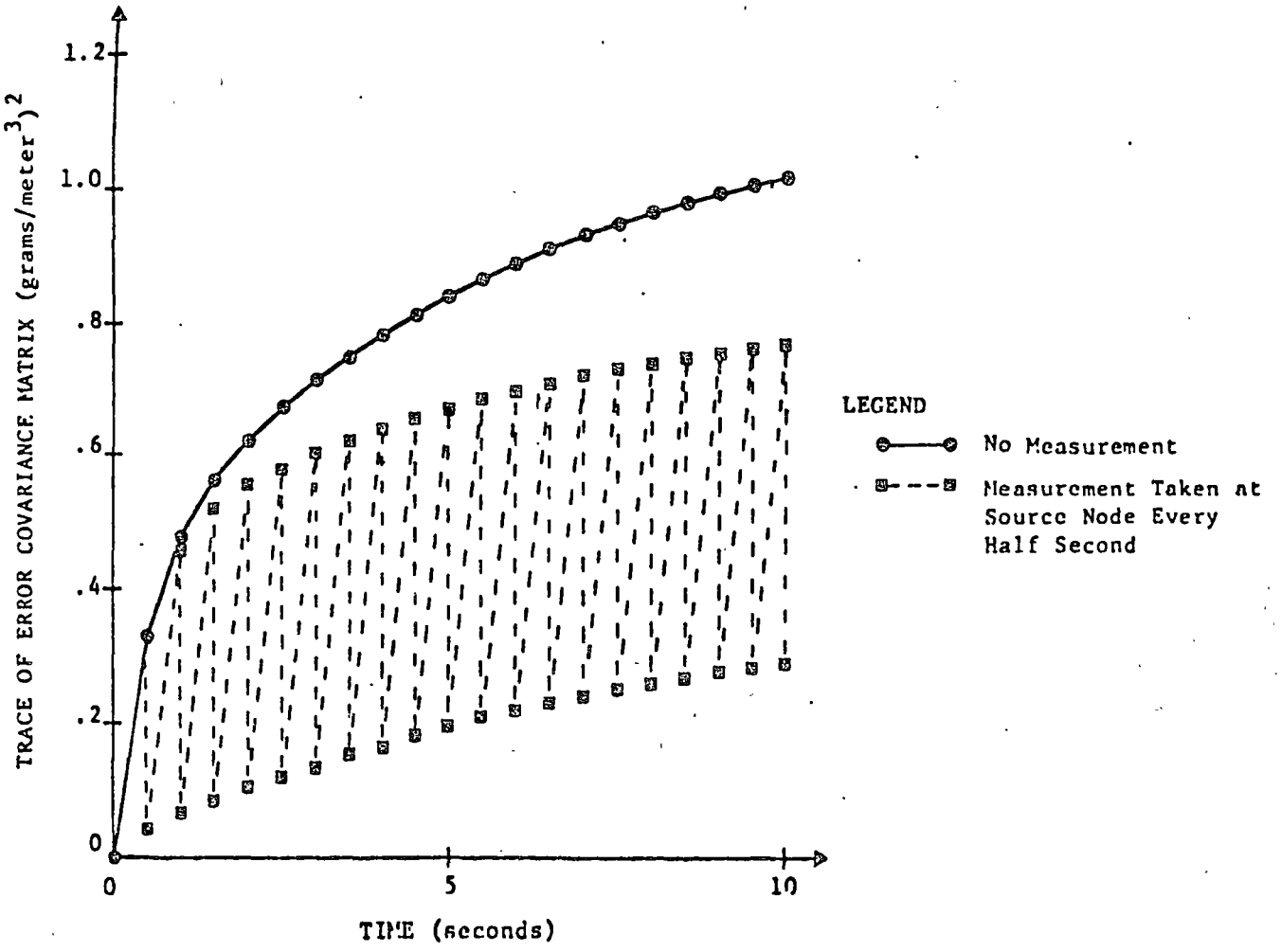


Figure 7 Temporal Effect of Sampling at Source Node Every Half Second

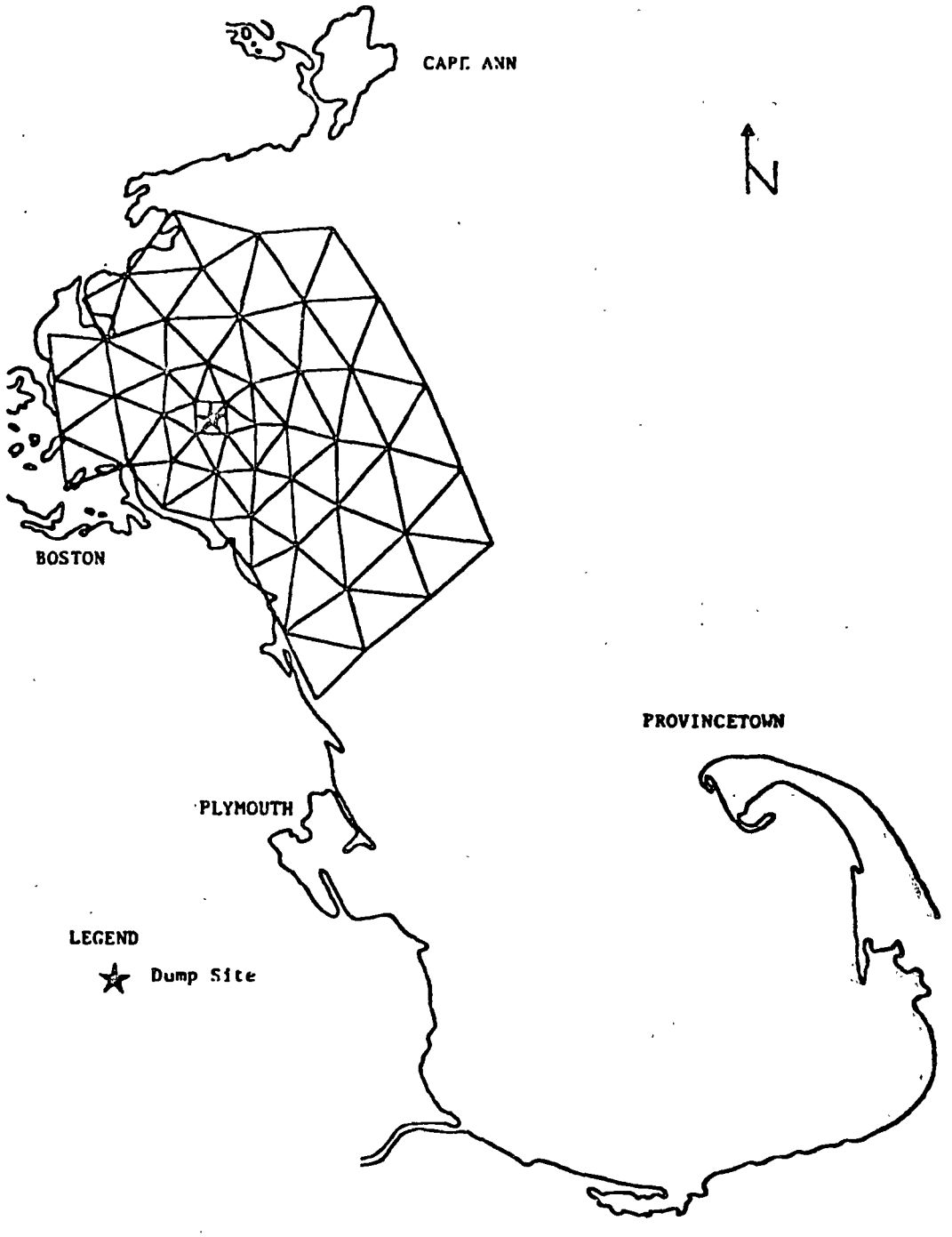


Figure 8 Finite Element Grid of Massachusetts Bay

the time period of the experiment.

In the NOMES experiment, approximately 1200 lbs. of sphalerite particles (mean diameter of approximately 5 microns; estimated 2.9×10^{15} particles) were dumped in a small area to simulate a point source discharge. Discharge began approximately one hour before low tide and lasted for approximately 10 minutes. Certain approximations are required in the mathematical modeling of the discharge. In the numerical simulation, it is necessary to take the duration of discharge equal to at least one time increment. (The time increment used in the dispersion model is 900 seconds.) In addition, although the discharge was essentially a point source, the source has to be distributed over a larger area (starred in Figure 8) in order to reduce the high concentration gradients which introduce numerical difficulties. In addition, the finite element grid is refined in the general discharge area for the same reason. Due to the spreading of the source over larger spatial and temporal scales, initial spreading is expected to be greater for the numerical results than in the actual experiment. With increasing time, this discrepancy should vanish.

The first order decay rate due to particle settling is obtained from Stoke's Law and assuming a uniform concentration profile over the water column depth as 3.3×10^{-5} meters/sec. The isotropic dispersion coefficient is chosen as 30 meters^2/sec (Pearce and Christodoulou [1975]).

Sensitivity of the dispersion model to uncertainty in the dispersion coefficient is addressed. Taking the standard deviation of the isotropic dispersion coefficient as 50% of the mean value (i.e., $\sigma_{E^{x,y}} = 15$ meters^2/sec), the effect on the predicted concentrations is shown in Figure 9 for a time of 12 hours after dump. Since dispersion is influenced by the concentration gradient, larger concentration uncertainties are expected in regions of steep concentration gradients. This is confirmed by the results shown.

Figure 10 illustrates the model sensitivity to a standard deviation of the decay rate equal to 50% of the mean value (i.e., $\sigma_{w_s} = 1.65 \times 10^{-5}$ meters/sec) at a time of 30 hours after dump. The model is observed to be less sensitive to a 50% of mean value standard deviation of the decay rate uncertainty than the dispersion uncertainty. The largest effect of the decay rate uncertainty is observed at the highest concentrations, as is expected.

The extent of application of the filtering algorithm for quantifying sampling effectiveness at the NOMES site was severely restricted by the high computational cost. For

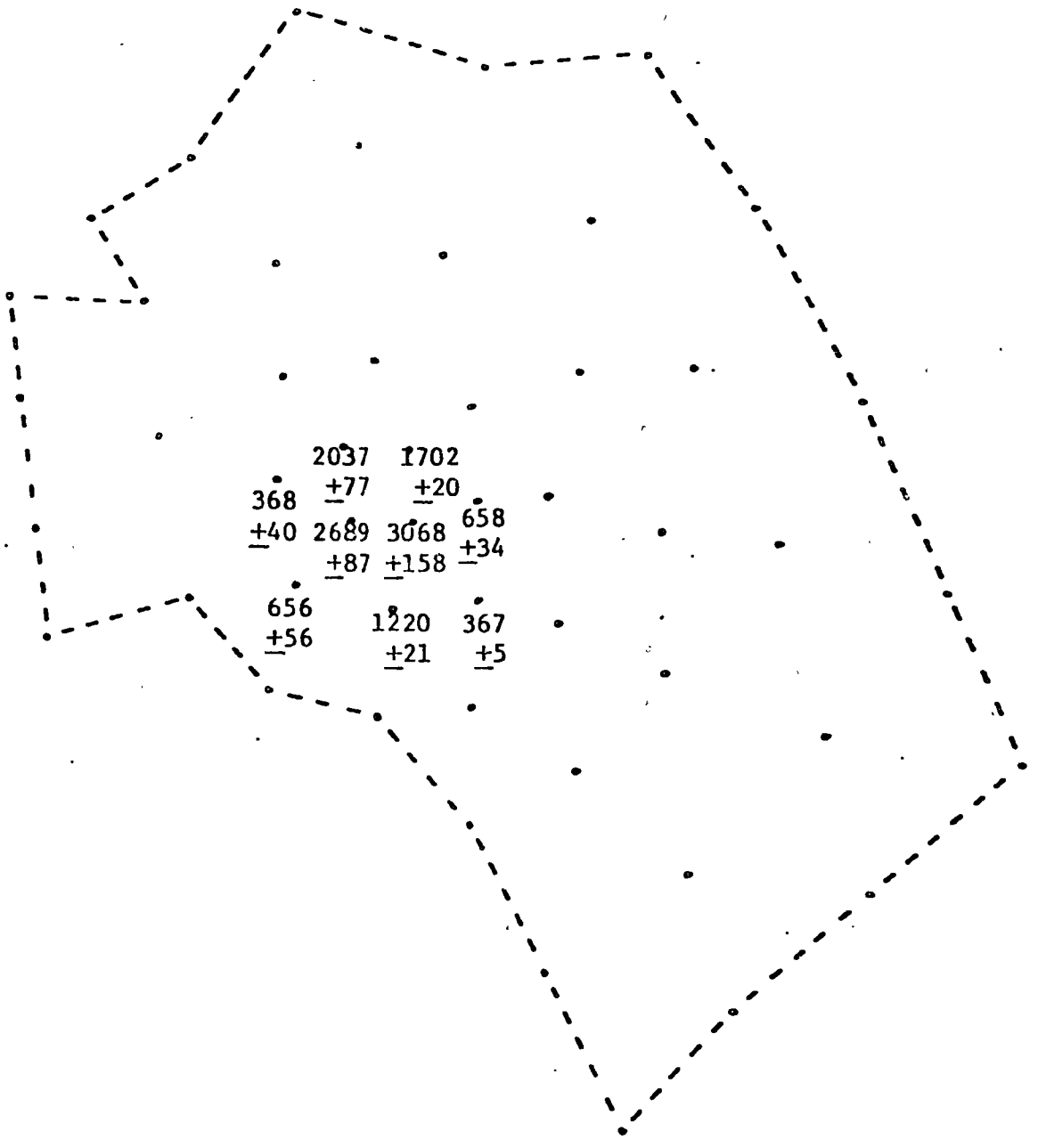


Figure 9 Mean Particle Concentrations and Standard Deviations of Modeling Uncertainty from a 15 meter^2/sec Standard Deviation of the Dispersion Coefficient at 12 Hours After Dump (particles/liter)

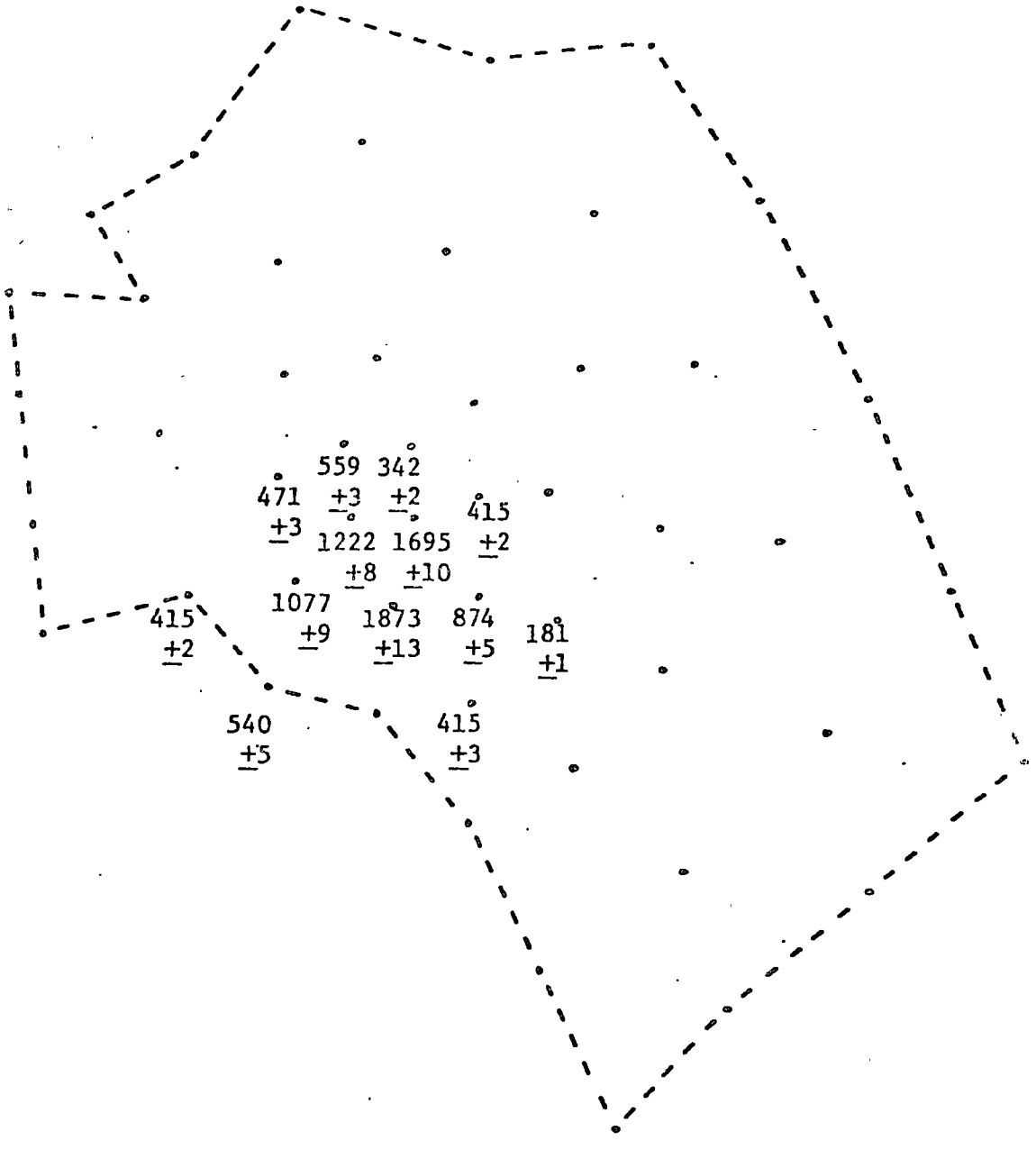


Figure 10 Mean Particle Concentrations and Standard Deviations of Modeling Uncertainty from a 1.65×10^{-5} meter/sec Standard Deviation of the Decay Rate at 30 Hours After Dump (particle/liter)

simulating the experiment for two full prototype days, the 50 state variable problem (see Figure 8) is required, but the computational problem is much too large for extensive simulation.

Given computational cost constraints, only one hypothetical sampling strategy was evaluated. Sampling was initiated at 9:00a.m. on the day following the dump (i.e., corresponding model time step 88). Samples were collected at the model source loading nodes every hour, until the completion of the sampling day at 4:00p.m. (i.e., corresponding to model time step 116). For purposes of presentation, the modeling uncertainty was computed from the uncertainty in the decay rate only. Figure 11 illustrates the effectiveness of the defined sampling strategy. In this figure, one observes the reduction in the system uncertainty due to the sampling effort. An interesting result is the very slow increase of the system uncertainty after the completion of sampling. Unfortunately, due to computational cost constraints, it was not possible to calculate the time duration after which no effect of the sampling would be felt.

CONCLUSIONS

Although a limited computer budget restricted the application to the NOMES experiment in Massachusetts Bay, these applications and extensive applications of the one-dimensional formulation have provided useful information on its computational costs and applicability.

The assessment of sampling effectiveness is made possible by filtering techniques. It allows the investigation of altered spatial and temporal frequency of sampling. However, the methodology does have limitations. One of the most critical is the requirement of the state-space representation of the system dynamics. Models are not generally developed in this form, due to the computational efficiency of other solution forms and the non-intuitive nature of the state-space form.

Although computation of the modeling uncertainty due to uncertainty in the dispersion coefficient, decay rate, velocity field, and source loading is presented, other uncertainty sources are not included. Uncertainty arises from assumptions made in the model formulation itself, which is difficult to quantify. For example, models are imperfect due to the assumptions of the applicability of Fickian diffusion and representation of naturally variant three-dimensional water bodies by lower dimensional models. Physical discretization

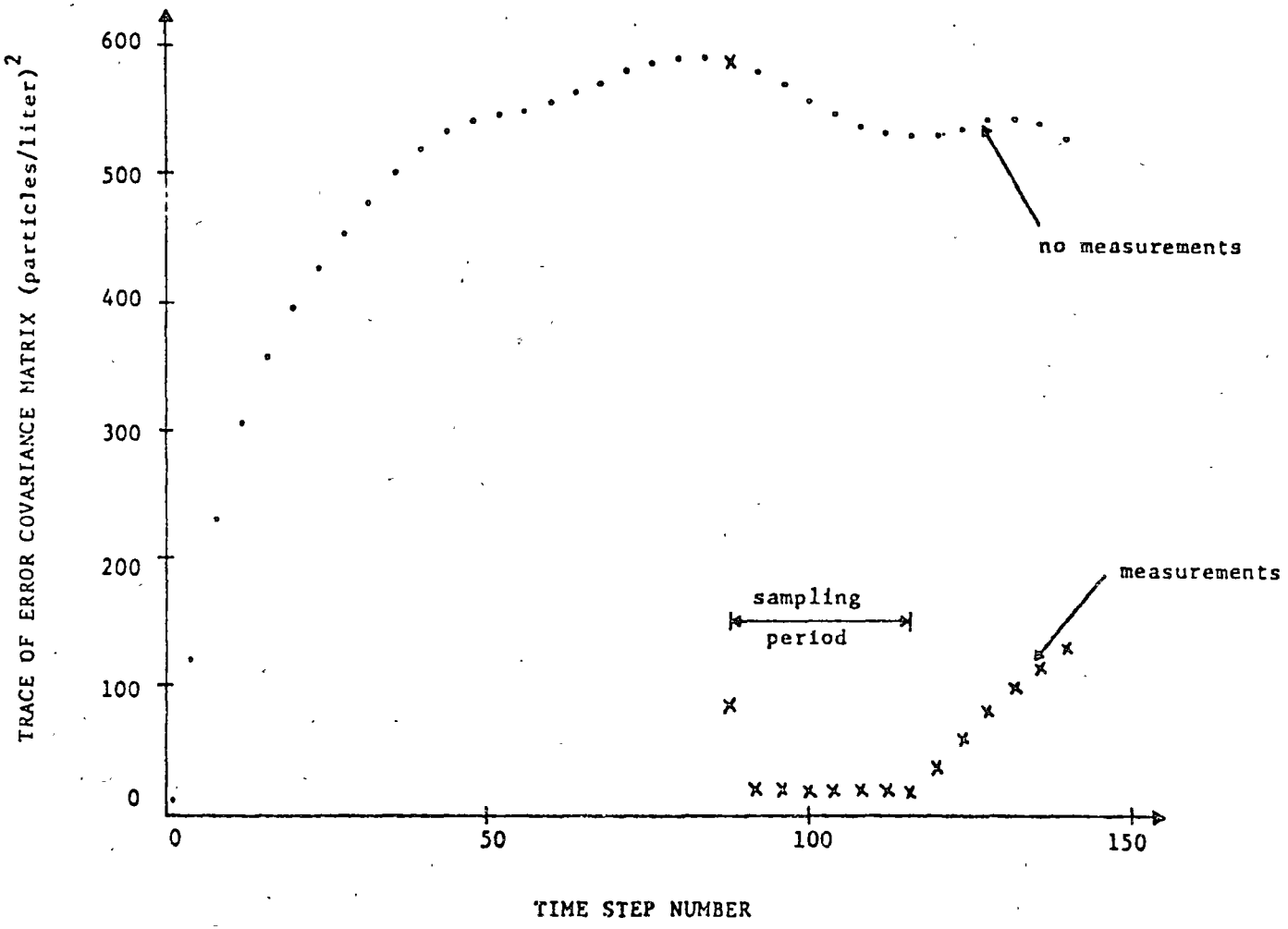


Figure 11 Temporal Effect of Hypothetical Sampling Strategy

of the continuous problem, both spatially and temporally, introduces numerical errors. By refining the grid and reducing the time increment, more accurate results are obtained, but at the expense of increased computational costs. In addition, round-off error can have significant effect on modeling results, but unfortunately has not received much attention.

The major reason for not including the effect on the modeling uncertainty of the above factors is the increased computational cost required. As is, calculation of the modeling uncertainty due to uncertainty in the parameters and inputs is costly for large systems. Calculation of the one-dimensional modeling uncertainty took roughly about 2 minutes of CPU time on an IBM 370 model 168 computer for a simulation of 21 nodes for 100 time steps. In comparison, calculation of the two-dimensional modeling uncertainty in application to Massachusetts Bay (neglecting uncertainty in the velocity field) took roughly 90 minutes of CPU time on the same computer for a simulation of 50 nodes for 144 time steps. Therefore, for large systems, the cost of computing the modeling uncertainty does become excessive; the additional cost of the Kalman filtering algorithm is insignificant.

An especially important conclusion of this study is the necessity to quantify the modeling uncertainty by a relatively detailed analysis. Many previous investigators (e.g., Moore (1973), Brewer and Moore (1974) and Pimentel (1975)) have obtained constant values of the modeling uncertainty based strictly upon subjective judgment. This practice is not advisable, as this work has shown the large spatial and temporal variability of the modeling uncertainty.

The tradeoff between computational cost and accuracy in quantifying the modeling uncertainty is evident. For simplistic one-dimensional modeling attempts, the relatively low computational cost justifies detailed modeling uncertainty analyses. The modeling uncertainty due to model assumptions, physical discretizations, round-off error, etc., should be addressed. On the other hand, difficulty in justifying the excessive computational costs of two-dimensional modeling of a Massachusetts Bay size problem exists. Although it is felt that the investment made in the simulation of a field experiment before it is actually performed will pay for itself in the higher return of information, the initial capital outlay for computational time may deter the use of such a methodology. Cheaper methods of calculating the modeling uncertainty are needed.

ACKNOWLEDGEMENTS

The research reported on in this paper was supported by the National Oceanographic and Atmospheric Administration Office of Sea Grant, Department of Commerce, through the M.I.T. Sea Grant Office.

REFERENCES

1. Brewer, J.W. and S.F. Moore, "Monitoring: An Environmental State Estimation Problem," ASME Journal of Dynamic Systems, Measurement and Control, pp. 363-365, 1974.

2. Dandy, G.C., "Design of Water Quality Sampling Systems for River Networks," Doctor of Philosophy Thesis, Department of Civil Engineering, Massachusetts Institute of Technology, June 1976.

3. DeGuida, R.N., "Design of Effective Sampling Programs for Verification of Coastal Dispersion," Master of Science Thesis, Department of Civil Engineering, Massachusetts Institute of Technology, May 1976.

4. Desalu, A.D., "Dynamic Air Quality Estimation in a Stochastic Dispersive Atmosphere," Doctor of Philosophy Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology, April 1974.

5. Gelb, A. (Editor), Applied Optimal Estimation, M.I.T. Press, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1974.

6. Jazwinski, A.H., Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.

7. Leimkuhler, W.F., "A Two-Dimensional Finite Element Dispersion Model," Engineer's Thesis, Department of Civil Engineering, Massachusetts Institute of Technology, September 1974.

8. Leimkuhler, W.F., Connor, J.J., Wang. J.D., Christodoulou, G.C., and Sundgren, S.L., "A Two-Dimensional Finite Element Dispersion Model," Proceedings of the Symposium on Modeling Techniques for Waterways, Harbors and Coastal Engineering, San Francisco, California, September 1975.

9. Lettenmaier, D.P., "Design of Monitoring Systems for Detection of Trends in Stream Quality." Technical Report No. 39, Charles W. Harris Hydraulics Laboratory, Seattle, Washington, August 1975.

10. Moore, S.F., "Estimation Theory Applications to Design of Water Quality Monitoring Systems," Journal of Hyd. Div., ASCE, HY5, pp. 815-831, May 1973.

11. Pearce, B.R. and Christodoulou, G.C., "Application of a Finite Element Dispersion Model for Coastal Waters," Proceedings XVI IAHR Congress, Sao Paulo, Brazil, July 1975.

12. Pearce, B.R., De Guida, R.N., Dandy, G.C. and Moore, S.F., "Sampling Network Design for Dispersion Verification," Proceedings of the Symposium on Modeling Techniques for Waterways, Harbors and Coastal Engineering, San Francisco, California, September 1975.

13. Pimentel, K.D., "Toward a Mathematical Theory of Environmental Monitoring: The Infrequent Sampling Problem," Ph.D. Thesis, University of California, Livermore, June 1975.

14. Schweppe, F.C., Uncertain Dynamic Systems, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1973.

15. Wang, J.D. and Connor, J.J., "Mathematical Modeling of Near Coastal Circulation," Technical Report No. 200, R.M. Parsons Laboratory for Water Resources and Hydrodynamics, Department of Civil Engineering, Massachusetts Institute of Technology, April 1975.

CHAPTER 3

3.1 FORMULATION

Beginning with the basic principles of conservation of mass and of force equilibrium,(Newton's second law), a formal mathematical model is developed for transient vertically integrated flow in the plane. The approach is somewhat similar to the works by Hansen [27], Reid and Bodine [59], Leendertse [38], Norton et. al. [49] and Pritchard [70]. We attempt to include all important steps of the development and to account for assumptions and their basis as much as possible. Where numerical parameters are needed in the constitutive equations, numbers or relationships based on experience are indicated. The model is thus intended to be truly predictive with the singular reservation that boundary conditions must be prescribed. The necessary boundary conditions for a well posed problem is also discussed.

3.2 THREE-DIMENSIONAL FLOW.

The mathematical formulation of the conservation of mass and momentum principles for three-dimensional flow has previously been derived in an eulerian framework using a cartesian x-y-z coordinate system, (see f.ex. [15]). The operation consists of balancing mass fluxes or forces for a small cube dx-dy-dz, (see Figure 3-1), and then taking the theoretical limit as the volume of the cube approaches zero. The result is

$$(3.2.1) \qquad \rho_{,t} + (\rho u)_{,x} + (\rho v)_{,y} + (\rho w)_{,z} = e$$

which states that the local rate of change of mass per volume, added to the net flux out, is equal to the rate of adding mass per volume, e. If there are no internal sources (henceforth we shall define a sink as a negative source and therefore only need to talk about sources), e is zero. ρ is the density; u,v,w are the velocity components in x,y,z directions and partial differentiation is written as a subscript comma followed by the independent variable. Equation (3.2.1) expresses a fundamental principle for any continuous one phase fluid.

The equilibrium of forces acting on the control volume is written for the x- and y-directions:

$$(3.2.2) \qquad (\rho u)_{,t} + (\rho u^2)_{,x} + (\rho uv)_{,y} + (\rho uw)_{,z} - \rho fv =$$

$$-p_{,x} + \tau_{xx,x} + \tau_{yx,y} + \tau_{zx,z} + \rho m_x$$

$$(3.2.3) \qquad (\rho v)_{,t} + (\rho uv)_{,x} + (\rho v^2)_{,y} + (\rho vw)_{,z} + \rho fu =$$

$$-p_{,y} + \tau_{xy,x} + \tau_{yy,y} + \tau_{zy,z} + \rho m_y$$

A rotating right handed coordinate system fixed on the earth with the z-axis vertical upwards is chosen. The equations (3.2.1) - (3.2.3) apply to the expected values of velocities and pressures which are considered to be stochastic processes. The τ's are therefore due to molecular viscosity and turbulent momentum transfer [15, 62]:

$$(3.2.4) \qquad \tau_{ij} = \overset{\nu}{\tau}_{ij} - \rho < u'_i u'_j > \qquad i,j = 1,2,3$$

where $< >$ signifies expected value of the argument, $\overset{\nu}{\tau}_{ij}$ is the viscous stress and u'_i is the turbulent velocity fluctuation in the i

direction. For convenience, here and in the following, frequent use of tensor notation will be made, the 1,2,3 directions being equivalent to x,y,z. The left hand sides of (3.2.2)and (3.2.3) represent the inertial forces on a unit volume and the right hand sides are the surface forces acting on the same volume plus internal sources of momentum m_x, m_y. In arriving at (3.2.2) and (3.2.3) is has been assumed that the vertical velocity w is small so that only $\rho f u$ and $\rho f v$ are retained from the fictitious coriolis force. f is the coriolis parameter = $2\,\omega_{earth} \cdot \sin\phi$, where ω_{earth} is the phase velocity of the earth's rotation and ϕ is the latitude (N) of the location.

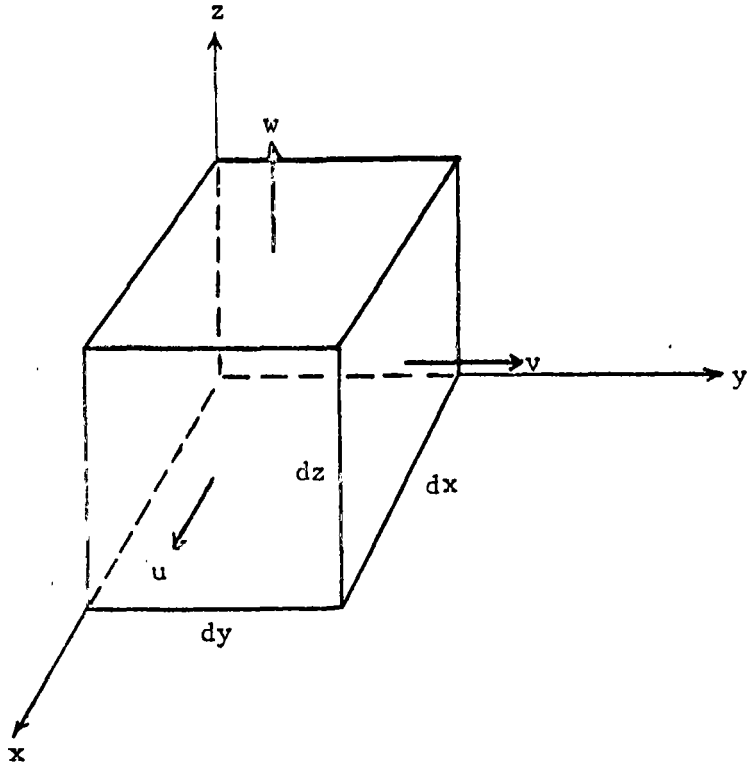


Figure 3-1. Infinitesimal control volume.

The isotropic normal stress in fluids is usually compressive and therefore denoted p for pressure (positive). The deviatoric stresses τ_{ij}, $i,j = 1,2,3$ are defined as usual, the first index denoting the normal direction of the face on which the stress acts and the second, the positive direction of the stress.

An order of magnitude comparison of the inertial terms in (3.2.2) and (3.2.3) is illustrative. Let $\hat{t}$, $\hat{\ell}$, $\hat{h}$, $\hat{u}$ and $\hat{w}$ be representative time, horizontal and vertical length and velocity scales. Scaling (3.2.2) then yields

$$\frac{\hat{u}}{\hat{t}} \sim \frac{\hat{u}^2}{\hat{\ell}} \sim \frac{\hat{u}\hat{w}}{\hat{h}} \sim \hat{f}\,\hat{u} \sim \hat{f}\,\hat{w}$$

where $\hat{f}\,\hat{w}$ is the so far ignored component of the coriolis force and $\hat{f} = f$ is equal to approximately 10^{-4}sec^{-1} at 40° latitude. In order to drop $\hat{f}\,\hat{w}$ and keep the remaining terms we must have

$$\hat{u} \gg \hat{w}\ ; \quad \frac{\hat{u}}{\hat{\ell}} \simeq \frac{\hat{w}}{\hat{h}} \Rightarrow \hat{\ell} \gg \hat{h}\ ; \quad \hat{t} \simeq \frac{\hat{\ell}}{\hat{u}}\ ; \quad \frac{\hat{u}}{\hat{\ell}} \simeq \hat{f}$$

For a typical coastal area $\hat{u} = 0(0.5\text{m/sec})$, $\hat{w} = 0(0.05\text{m/sec})$ $\hat{\ell} = 0(10^3\text{m})$, $\hat{h} \simeq 0(100\text{m})$ giving a corresponding time scale $\hat{t} = 0(2.10^3\text{sec}) = 0(0.6\text{ hr})$, indeed in agreement with the above scaling relations.

Vertical equilibrium requires

$$(3.2.5) \qquad (\rho w)_{,t} + (\rho uw)_{,x} + (\rho vw)_{,y} + (\rho w^2)_{,z} + 2\rho\omega_x v - 2\rho\omega_y u =$$

$$- p_{,z} - \rho g + \tau_{xz,x} + \tau_{yz,y} + \tau_{zz,z}$$

where ω_x and ω_y are the x and y components of the earth's rotation.

Scaling of this equation leaves only the pressure, gravity and normal stress terms as significant. Again τ_{zz} is related to molecular viscosity and the vertical velocity fluctuations, hence it can be neglected in comparison with ρg and we finally obtain the hydrostatic pressure condition.

$$p_{,z} = -\rho g \tag{3.2.6}$$

Along the boundaries, special conditions apply. Thus the fact that the free surface is a material interface is expressed as the kinematical condition

$$\frac{D}{Dt}(\eta - z)\Big|_{z=\eta} = \left[\frac{\partial \eta}{\partial t} + u\frac{\partial \eta}{\partial x} + v\frac{\partial \eta}{\partial y} - w\right]_{z=\eta} = 0 \tag{3.2.7}$$

where evaporation and rainfall are neglected. $\eta = \eta(x,y,t)$ is the surface elevation, Figure 3-2, and $\frac{D}{Dt}$ is the total or particle derivative. At the bottom $z = -h(x,y)$, which is assumed fixed and impermeable, the similar condition is

$$\frac{D}{Dt}(z+h)\Big|_{z=-h} = \left[u\frac{\partial h}{\partial x} + v\frac{\partial h}{\partial y} + w\right]_{z=-h} = 0 \tag{3.2.8}$$

For lateral boundaries, which are assumed vertical, (see Figure 3-3) the flow must be continuous, implying

$$u_i n_i\Big|_-^+ = 0 \qquad i = 1,2 \tag{3.2.9}$$

$$\ell_{ijk} u_i n_j\Big|_-^+ = 0 \qquad i,j,k = 1,2,3 \tag{3.2.10}$$

where ℓ_{ijk} is the permutation tensor, $\ell_{123} = \ell_{312} = \ell_{231} = 1$ $\ell_{132} = \ell_{213} = \ell_{321} = -1$ and all other elements are zero.

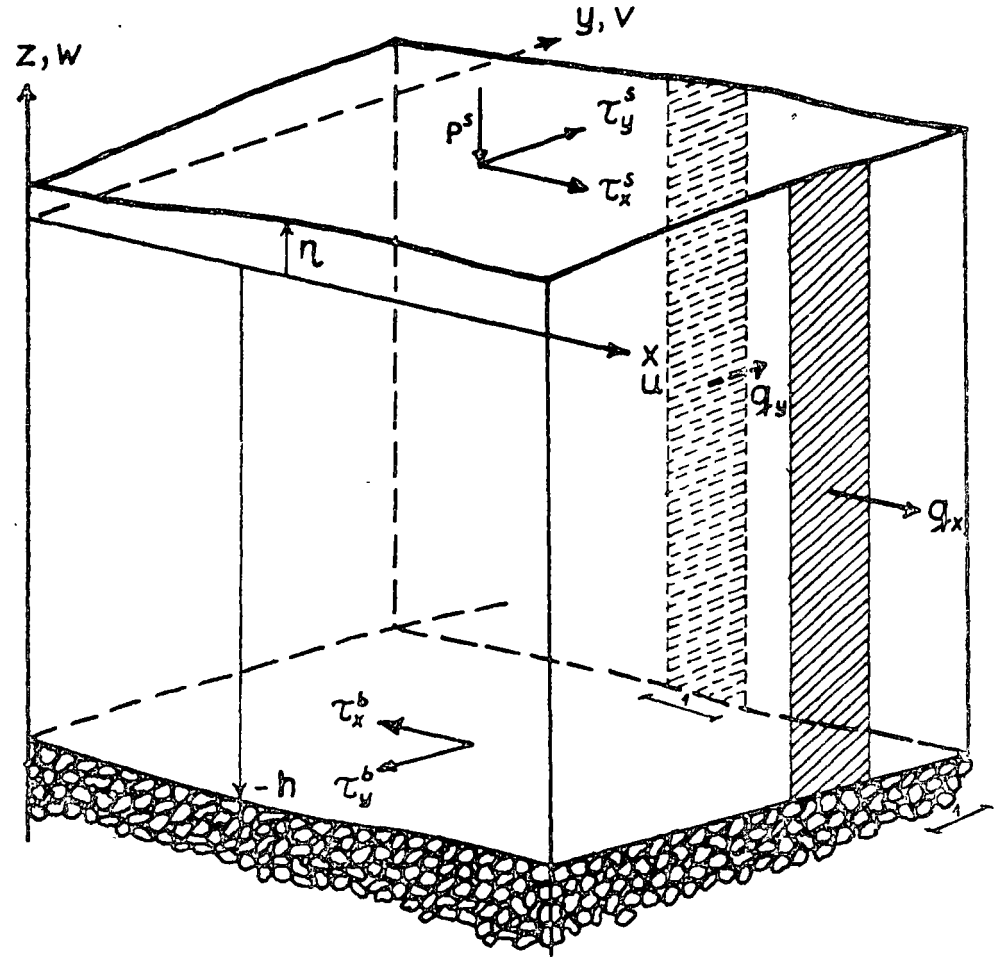


Figure. 3-2. Definition sketch.

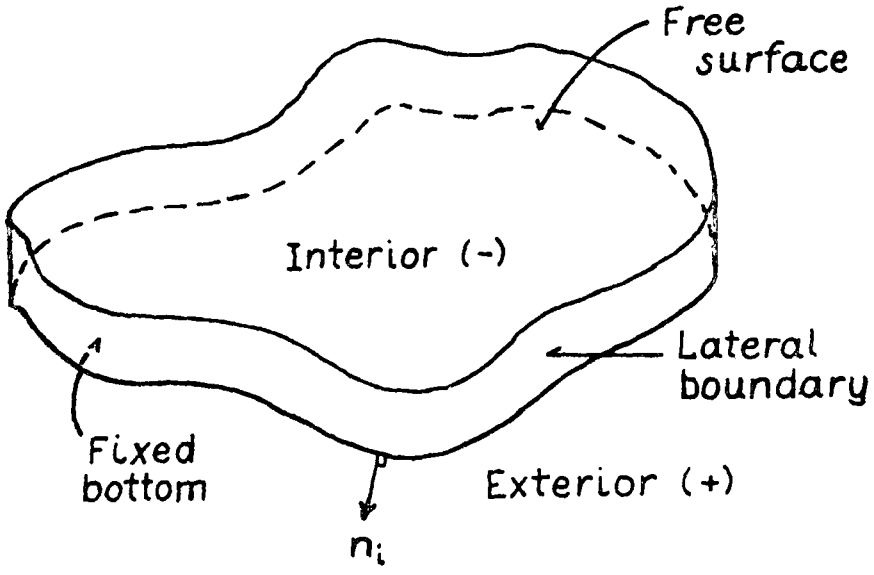


Figure 3-3. Domain terminology.

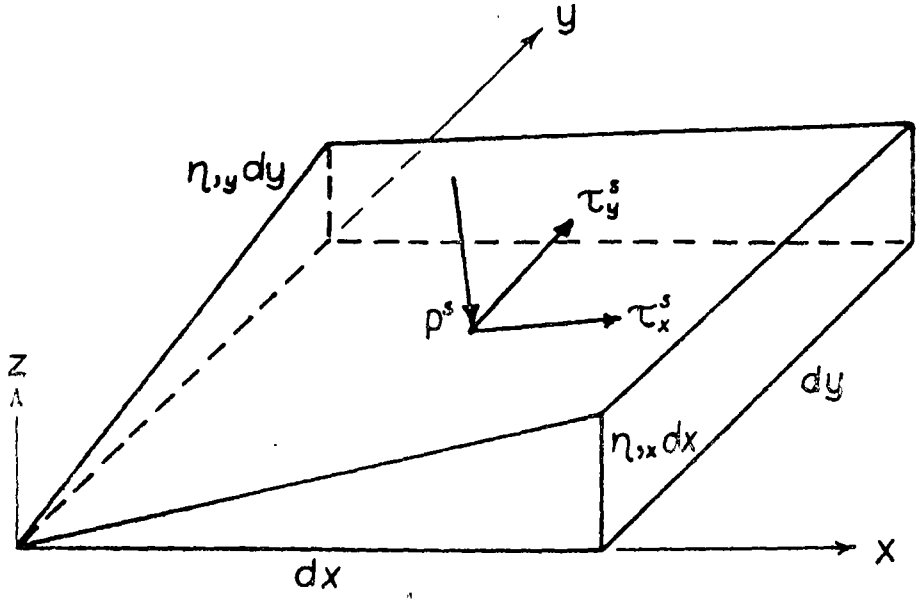


Figure 3-4. Surface force notation.

n_i is a unit vector perpendicular to the boundary directed out of the area of interest. + and − are points just outside and inside of the boundary. Equations (3.2.9), (3.2.10) express that normal and tangential velocities just outside and inside the boundary must be equal.

Dynamic equilibrium must also be satisfied on the boundaries, see Figure 3-4. Projecting the forces for a surface element on x,y and z-directions result in

$$(3.2.11)\qquad (\tau_x^s + p^s \eta_{,x}) = \left[p\eta_{,x} - \tau_{xx}\eta_{,x} - \tau_{yx}\eta_{,y} - \tau_{zx}\right]_{z=\eta}$$

$$(3.2.12)\qquad (\tau_y^s + p^s \eta_{,y}) = \left[p\eta_{,y} - \tau_{xy}\eta_{,x} - \tau_{yy}\eta_{,y} - \tau_{zy}\right]_{z=\eta}$$

$$(3.2.13)\qquad -p^s + \tau_x^s \eta_{,x} + \tau_y^s \eta_{,y} = \left[-p - \tau_{xz}\eta_{,x} - \tau_{yz}\eta_{,y} + \tau_{zz}\right]_{z=\eta}$$

and similarly for the bottom $(z = -h(x,y))$.

$$(3.2.14)\qquad \tau_x^b - p^b h_{,x} = \left[-(p - \tau_{xx}) h_{,x} + \tau_{yx} h_{,y} + \tau_{zx}\right]_{z=-h}$$

$$(3.2.15)\qquad \tau_y^b - p^b h_{,y} = \left[-\tau_{xy} h_{,x} - (p - \tau_{yy}) h_{,y} + \tau_{zy}\right]_{z=-h}$$

$$(3.2.16)\qquad -p^b - \tau_x^s h_{,x} - \tau_y^s h_{,y} = \left[-p + \tau_{xz} h_{,x} + \tau_{yz} h_{,y} + \tau_{zz}\right]_{z=-h}$$

On lateral boundaries, continuity of the stresses is again required.

$$(3.2.17)\qquad [\text{normal stress}]_-^+ = 0$$

$$(3.2.18)\qquad [\text{tangential stress}]_-^+ = 0$$

In case the fluid is considered inviscid (3.2.10) and (3.2.18)

must be relaxed and only (3.2.9) and (3.2.17) enforced. Finally, for a well defined problem, the initial flow field u_i^{in} must be known:

$$(3.2.19) \qquad u_i^{in} = u_{o_i} \quad i = 1,2,3 \quad x,y,z \in \Omega \quad t = 0$$

For an arbitrary geometry, the problem as formulated is not easily solved. Numerical solutions are stymied by excessive computer requirements and lack of information on the proper boundary conditions. In coastal areas that are well mixed through the water column, a significant simplification is achieved by eliminating the explicit dependence on the vertical coordinate. This process is described in the following section.

3.3 VERTICAL INTEGRATION

In shallow water bodies, the flow variation through the depth is often less significant. In such cases, vertically integrated equations and variables may adequately describe the situation. This approach yields estimates for the transport through any cross section, however, detailed information on the velocity structure is lost. In the following, the water density is assumed constant in the z direction, i.e. $\rho = \rho(x,y,t)$. This and the assumption of relatively small vertical velocities and accelerations are normally implied by the definition shallow.

The development of a boundary layer from an applied wind stress on the surface is dependent on the magnitude of the vertical turbulent momentum transfer. Several investigations have found the vertical eddy viscosity falling in the range $E \sim 1 - 200$ cm^2/sec. If the time scale of 1 hour is retained, a notion of the meaning of the expression

shallow in connection with wind driven circulation is obtained from the expression

$$(3.3.1) \qquad h_w = \sqrt{E \cdot \hat{t}} \sim 1 - 10 \text{ m}$$

The use of vertically integrated quantities to predict dispersion of a tracer in this situation is clearly less satisfactory because of the highly non-uniform velocity profile, although the total transport still may be well predicted. In two and multi-layered models, some improvement on this point can be expected. For tidal flow, the driving force which is the hydrodynamic pressure, acts over the entire depth, and vertically integrated values are expected to be representative for the local velocities also, except close to the bottom. Finally, neutrally buoyant stream flows entering or leaving the area are well suited for an integrated treatment since those flows generally are well mixed. Again, in-or outflows with a density difference are better simulated in multilayer models.

The governing equations are integrated over the total depth to eliminate the z-dependence. Beginning with Equation (3.2.1), we formally write

$$(3.3.2) \qquad \int_{-h}^{\eta} \rho_{,t}\, dz + \int_{-h}^{\eta} \rho(u)_{,x}\, dz + \int_{-h}^{\eta} (\rho v)_{,y}\, dz + \int_{-h}^{\eta} (\rho w)_{,z} dz = \int_{-h}^{\eta} e\, dz$$

Making use of Leibnitz's rule [29] we may change the order of integration and differentiation to obtain

$$(3.3.3) \qquad \frac{\partial}{\partial t}\int_{-h}^{\eta} \rho dz - \rho\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x}\int_{-h}^{\eta} \rho u dz - \rho u\Big|_{\eta}\frac{\partial \eta}{\partial x} + \rho u\Big|_{-h}\frac{\partial(-h)}{\partial x}$$

$$+ \frac{\partial}{\partial y}\int_{-h}^{\eta} \rho v dz - \rho v\Big|_{\eta}\frac{\partial \eta}{\partial y} + \rho v\Big|_{-h}\frac{\partial(-h)}{\partial y} + \rho w\Big|_{\eta} - \rho w\Big|_{-h} = \rho q_I$$

Finally applying the kinematic conditions on the surface and bottom results in

$$(3.3.4) \qquad (\rho H)_{,t} + (\rho q_x)_{,x} + (\rho q_y)_{,y} = \rho q_I$$

where we have introduced the integrated variables, total depth:

$$(3.3.5) \qquad H = \int_{-h}^{\eta} dz = h + \eta$$

discharges in x- and y-directions per unit width:

$$(3.3.6) \qquad q_x \equiv \int_{-h}^{\eta} u \, dz$$

$$(3.3.7) \qquad q_y \equiv \int_{-h}^{\eta} v \, dz$$

If we let q_I represent the net rate of volume addition per unit horizontal area, this result (3.3.4) is generally valid for any type of flow, including situations with permeable bottom and evaporation or precipitation at the free surface. The primary objective for including q_I is however to make possible modeling of internal sources such as the discharge from a diffuser pipe.

The integration of the momentum equations, (3.2.2) - (3.2.3) proceeds analogously:

(3.3.8)
$$\int_{-h}^{\eta} (\rho u)_{,t}\,dz + \int_{-h}^{\eta} \{(\rho u^2)_{,x} + (\rho uv)_{,y}\}\,dz + \int_{-h}^{\eta} (\rho uw)_{,z}\,dz$$

$$-\int_{-h}^{\eta} \rho f v\,dz + \int_{-h}^{\eta} p_{,x}\,dz - \int_{-h}^{\eta} (\tau_{xx,x} + \tau_{yx,y})\,dz - \int_{-h}^{\eta} \tau_{zx,z}\,dz$$

$$-\int_{-h}^{\eta} \rho m_x\,dz =$$

$$\frac{\partial}{\partial t}\int_{-h}^{\eta} (\rho u)\,dz - (\rho u)\Big|_{z=\eta}\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x}\int_{-h}^{\eta} (\rho u^2)\,dz - (\rho u^2)\Big|_{\eta}\frac{\partial \eta}{\partial x}$$

$$+ (\rho u^2)\Big|_{-h}\frac{\partial(-h)}{\partial x} + \frac{\partial}{\partial y}\int_{-h}^{\eta} (\rho uv)\,dz - (\rho uv)\Big|_{\eta}\frac{\partial \eta}{\partial y} + (\rho uv)\Big|_{-h}\frac{\partial(-h)}{\partial y}$$

$$+ (\rho uw)\Big|_{\eta} - (\rho uw)\Big|_{-h} - \rho f q_y + \frac{\partial}{\partial x}\int_{-h}^{\eta} p\,dz - p\Big|_{\eta}\frac{\partial \eta}{\partial x}$$

$$+ p\Big|_{-h}\frac{\partial(-h)}{\partial x} - \frac{\partial}{\partial x}\int_{-h}^{\eta} \tau_{xx}\,dz + \tau_{xx}\Big|_{\eta}\frac{\partial \eta}{\partial x} - \tau_{xx}\Big|_{-h}\frac{\partial(-h)}{\partial x}$$

$$- \frac{\partial}{\partial y}\int_{-h}^{\eta} \tau_{yx}\,dz + \tau_{yx}\Big|_{\eta}\frac{\partial \eta}{\partial y} - \tau_{yx}\Big|_{-h}\frac{\partial(-h)}{\partial y} - (\tau_{zx})\Big|_{\eta} + (\tau_{zx})\Big|_{-h}$$

$$-\rho \bar{m}_x =$$

$$\frac{\partial}{\partial t}(\rho q_x) + \frac{\partial}{\partial x}\int_{-h}^{\eta} (\rho u^2)\,dz + \frac{\partial}{\partial y}\int_{-h}^{\eta} (\rho uv)\,dz - \rho f q_y + \frac{\partial}{\partial x}\int_{-h}^{\eta} p\,dz$$

$$- \frac{\partial}{\partial x}\rho F'_{xx} - \frac{\partial}{\partial y}\rho F'_{yx} - \tau_x^s + \tau_x^b - \rho \bar{m}_x - p^s\frac{\partial H}{\partial x} - p^b\frac{\partial h}{\partial x} = 0$$

in which we have defined

$$\rho F_{xx}' \equiv \int_{-h}^{\eta} \tau_{xx}\, dz \tag{3.3.9}$$

$$\rho F_{yx}' \equiv \int_{-h}^{\eta} \tau_{yx}\, dz \tag{3.3.10}$$

$$\bar{m}_x \equiv \int_{-h}^{\eta} m_x\, dz \tag{3.3.11}$$

For computational reasons it is more convenient to work with the pressure in excess of hydrostatic pressure corresponding to the water level at datum and rest.

The density may be written as a mean value plus a deviation

$$\rho(x,y,t) = \rho_o + \Delta\rho(x,y,t) \tag{3.3.12}$$

and assuming the instantaneous local deviation is small compared to the mean

$$\Delta\rho \ll \rho_o \tag{3.3.13}$$

Boussinesque's approximation [57] is introduced whereby the density in all terms is replaced by the constant mean density ρ_o. This is a reasonable simplification provided the real density is used in the pressure term which now takes the form

$$\rho_o F_p = \int_{-h}^{\eta} p\, dz - \frac{1}{2}\rho_o\, g\, h^2 \tag{3.3.14}$$

$$\approx \rho_o g\, h\eta + \frac{1}{2}\rho_o\, g\, \eta^2 + \frac{1}{2}\Delta\rho\, g\, H^2 + p^s H$$

With these definitions and approximations, the final form of the equilibrium equation (3.3.8) becomes

$$(3.3.15)\qquad \frac{\partial}{\partial t}(q_x) + \frac{\partial}{\partial x}\int_{-h}^{\eta} u^2 dz + \frac{\partial}{\partial y}\int_{-h}^{\eta} uv\, dz - fq_y + \frac{\partial}{\partial x}(F_p - F_{xx}') - \frac{\partial}{\partial y}F_{yx}' + \frac{\tau_x^s - \tau_x^b}{\rho_o} - \bar{m}_x - \frac{p^s}{\rho_o}\frac{\partial H}{\partial x} - g\eta\frac{\partial h}{\partial x} - \frac{\Delta\rho}{\rho_o} gH\frac{\partial h}{\partial x} = 0$$

By complete analogy, the force balance in y-direction gives

$$(3.3.16)\qquad \frac{\partial}{\partial t}(q_y) + \frac{\partial}{\partial x}\int_{-h}^{\eta} uv\, dz + \frac{\partial}{\partial y}\int_{-h}^{\eta} v^2 dz + fq_x - \frac{\partial}{\partial x}F_{xy}' + \frac{\partial}{\partial y}(F_p - F_{yy}') + \frac{\tau_y^s - \tau_y^b}{\rho_o} - \bar{m}_y - \frac{p^s}{\rho_o}\frac{\partial H}{\partial x} - \frac{\Delta\rho}{\rho_o} g H\frac{\partial h}{\partial y} - g\eta\frac{\partial h}{\partial x} = 0$$

with the corresponding definitions:

$$(3.3.17)\qquad \rho F_{xy}' \equiv \rho F_{yx}' \equiv \int_{-h}^{\eta} \tau_{yx}\, dz$$

$$(3.3.18)\qquad \rho F_{yy}' \equiv \int_{-h}^{\eta} \tau_{yy}\, dz$$

$$(3.3.19)\qquad \bar{m}_y \equiv \int_{-h}^{\eta} m_y\, dz$$

The number of unknowns still exceeds the number of equations for our problem. To overcome this hurdle, the currently most successfully used empirical relations for bottom and surface friction are reviewed in order to establish a set of constitutive equations. Previous modeling has shown that a quadratic, (in mean velocity), bottom friction law in all cases adequately represents the damping due to the shear at the bottom. Several similar empirical expressions, Manning, Chezy, and Darcy-Weisbach equa-

tions [15], were originally derived from measurements of steady flow in channels or pipes; but have been modified for two-dimensional unsteady circulation. The quoted relationships are the most widely used and relates shear stress to discharge per unit width as follows:

$$\tau_x^b = C_f \rho (q_x^2 + q_y^2)^{1/2} \frac{q_x}{H^2} \tag{3.3.20}$$

$$\tau_y^b = C_f \rho (q_x^2 + q_y^2)^{1/2} \frac{q_y}{H^2} \tag{3.3.21}$$

where

$$C_f = \begin{cases} \frac{1}{8} f_{DW} & \text{Darcy-Weisbach} \\ \frac{g}{C^2} & \text{Chezy} \\ \frac{n^2}{H^{1/3}} g & \text{Manning} \end{cases} \tag{3.3.22}$$

Values of Manning's n are only known for fully developed rough turbulent flow, which fortunately is the normal case in coastal areas, as the Reynolds number $R = \frac{u \cdot H}{\nu} = \frac{1 \cdot 10}{10^{-6}} = 10^7$ and the relative roughness $\frac{k_s}{H} \simeq 0.01 - 0.1$. For fixed roughness, the friction factor C_f is therefore inversely proportional to $H^{1/3}$. Normal values of n range 0.025 - 0.040. The values of C_f for some n and depth values are given in Table 3-1.

In other flow regimes, the use of a Moody diagram to find f_{DW} is the best approach. Choosing as an example $C_f = 0.005$ and a velocity of 1 m/sec gives a shear stress of 5 N/m^2 which is considered as a large bottom friction.

Bottom roughness k_s [m]	n [sec $\cdot m^{-1/3}$] \ H [m]	1	2	5	10	20	30	40	50	100
Stones 0.07	0.025	0.0061	0.0049	0.0036	0.0028	0.0023	0.0020	0.0018	0.0017	0.0013
Small rocks 0.20	0.030	0.0088	0.0070	0.0052	0.0041	0.0033	0.0028	0.0026	0.0024	0.0019
Dunes 0.50	0.035	-	0.0095	0.0070	0.0056	0.0044	0.0039	0.0035	0.0033	0.0026
1.10	0.040	-	-	0.0092	0.0073	0.0058	0.0051	0.0046	0.0043	0.0034

TABLE 3-1: Values of C_f

The wind stress on the surface is more complicated to handle because the water surface is deformable so that waves form, and also the length scale of the turbulent wind field is so large that the wind stress is highly variable in time and space [17, 25].

Several investigators have derived expressions for the average wind stress from measurements in the field [5, 16, 28, 74, 77, 79]. If the shear stress is related to the wind speed as follows,

$$\tau^s = \rho_{air} C_D U_{10}^2 \qquad (3.3.23)$$

where ρ_{air} is the air density, ($\sim$1.2 kg/m^3) and U_{10} is the wind speed at 10 m above the surface, then the wind drag coefficient C_D has been found to vary from approximately 0.001 and up according to the following relations:

$$(3.3.24)\qquad C_D = \begin{cases} 1.25/U_{10}^{1/5} \cdot 10^{-3} & U_{10} \leq 1 \text{ m/sec} \\ 0.5 \cdot U_{10}^{1/2} \cdot 10^{-3} & 1 < U_{10} \leq 15 \text{ m/sec} \quad \text{Ref.[79]} \\ 2.6 \cdot 10^{-3} & U_{10} \geq 15 \text{ m/sec.} \end{cases}$$

$$(3.3.25)\qquad C_D = \begin{cases} 1.0 \cdot 10^{-3} & U_{10} \leq 5.6 \text{ m/sec} \quad \text{Ref.[74]} \\ 1.0 + 1.9\left(1 - \frac{5.6}{U_{10}}\right)^2 \cdot 10^{-3} & U_{10} > 5.6 \text{ m/sec.} \end{cases}$$

$$(3.3.26)\qquad C_D = 0.00228 + (1.0 - 7.0/U_{10})^2\ 0.00263 \qquad 20 \leq U_{10} \leq 40 \text{ m/sec} \qquad \text{Ref.[77]}$$

$$(3.3.27)\qquad C_D = \begin{cases} 0.577 \cdot 10^{-3} & U_{10} \leq 4.9 \text{ m/sec} \quad \text{Ref.[28]} \\ (-0.125 + 0.1427\ U_{10}) \cdot 10^{-3} & 4.9 \leq U_{10} \leq 19.2 \text{ m/sec.} \\ 2.62 & 19.2 < U_{10} \text{ m/sec} \end{cases}$$

The values given in the referenced papers are plotted in Figure 3-5. The data in [5, 74, 77] were for ponds or lakes, and [16,28,79] used measurements on the open ocean. There is a significant scatter of the data and hence of the curves used to fit the data points. Wu's relationships based on ocean data seem to give the best overall fit. Unfortunately, there are two discontinuities in the suggested relation for C_D, (3.3.24) which physically does not seem reasonable although some justification is attempted [79]. Considering the spread of the curves with a factor of 2 difference between results, it is tempting to fit one straight line

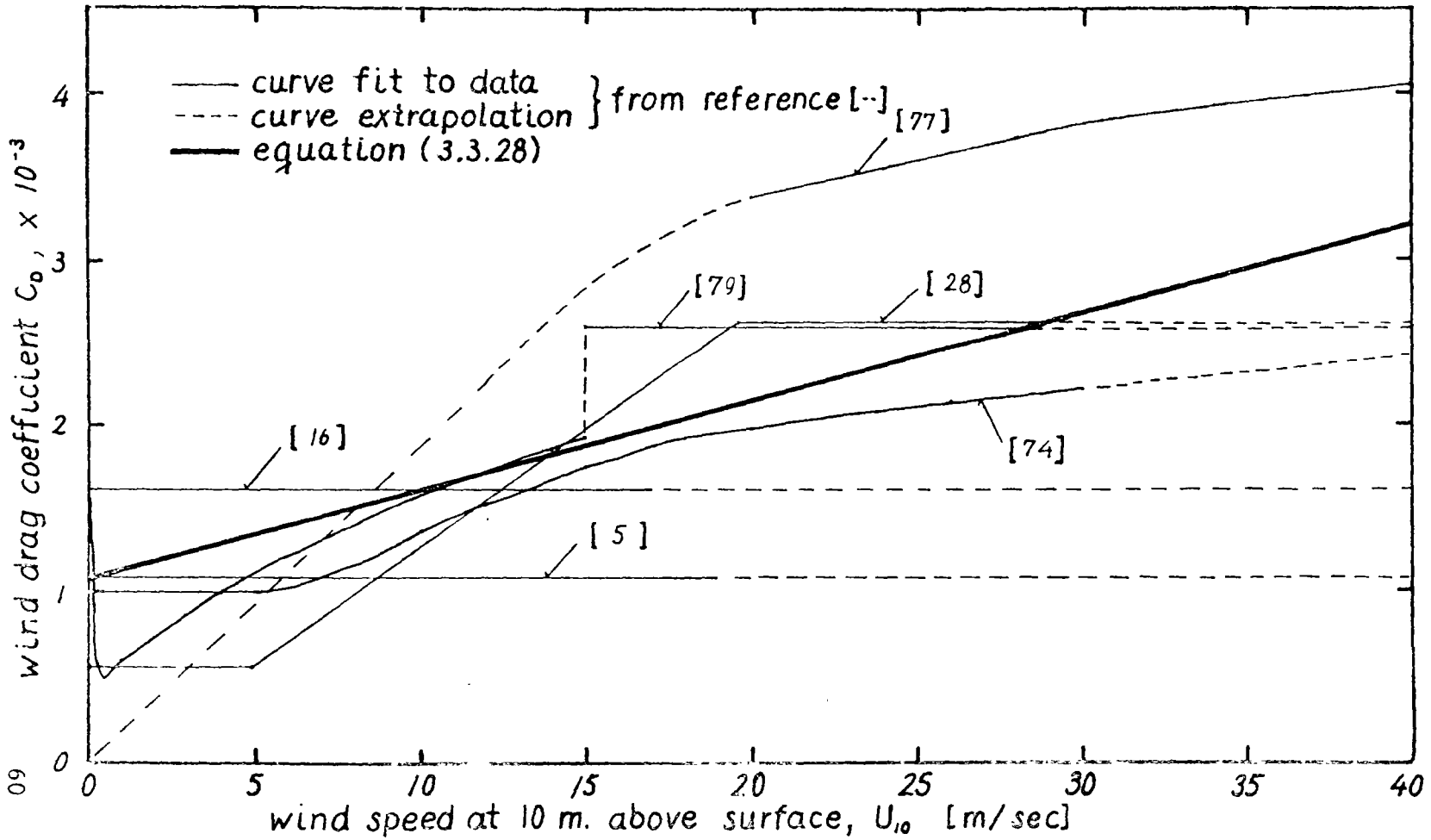


Figure 3-5. Wind drag coefficient, C_D, vs. wind speed, U_{10}.

relation as shown, with the equation

$$(3.3.28) \qquad C_D = \{1.1 + 0.0536 \cdot U_{10}\} \cdot 10^{-3} \qquad U_{10} \text{ in [m/sec]}$$

For U_{10} = 10 m/sec, the drag coefficient is $1.64 \cdot 10^{-3}$ and the predicted shear stress $\tau^s = 0.2$ N/m^2 which is somewhat larger than the ~0.1 N/m^2 normally measured in Massachusetts Bay for similar winds.† For wind speeds ranging from 0 - 30 m/sec, we can conclude, the present state of the art only allows us to predict the applied wind stress to within a factor of 2. However, considering the complexity of this problem, such an error seems tolerable.

Finally, the origin and significance of the internal stress terms τ_{xx}, $\tau_{xy} = \tau_{yx}$, τ_{yy} are investigated. To close the formulation we also try to express these terms as functions of the integrated flow variables by means of an eddy viscosity coefficient matrix. The approach, in many ways similar to the closure of turbulent flow problems [62], is admittedly based on a physically very loose foundation; but does yield an attractive structure reflecting many of the expected real effects, viz dissipation, and diffusion of momentum. The vertically integrated approach is only valid when the internal stresses are relatively small, so an exact representation of these terms is assumed to be of minor importance. All this trouble is directly caused by averaging the convective acceleration terms. However, the real root of the problem is the use of eulerian rather than lagrangian description (in the latter, the observer follows a particle and

† In the range 0 - 10 m/sec, Equation (3.3.28) agrees well with some new results by Parker and Pearce [55].

the convective terms do not appear). We have to live with the eulerian viewpoint in which the observer is fixed in space and propose taking a closer look at the implication of ensemble averaging and vertically integrating an instantaneous velocity product. Without loss of generality we may write the instantaneous local velocity components U,V as

$$U \equiv u + u' \equiv (\bar{u} + \bar{u}') + (u'' + u''') \tag{3.3.29}$$

$$V \equiv v + v' \equiv (\bar{v} + \bar{v}') + (v'' + v''') \tag{3.3.30}$$

where u,v are ensemble averages (assuming the flow field is basically random); u', v' are random fluctuations whose ensemble means per definition are zero; $\bar{u}$, $\bar{v}$ are the vertical average values of u, v; $\bar{u}'$, $\bar{v}'$ are vertical average values of u', v'; u'', v'' are vertical deviations of u,v from $\bar{u}$, $\bar{v}$; and finally, u''', v''' are vertical deviations of u', v' from $\bar{u}'$, $\bar{v}'$. The significance of each of the various components is shown in Fig. 3-6.

The product U·V is now written out in terms of its components

$$\begin{aligned} U \cdot V &= (\bar{u} + \bar{u}' + u'' + u''')(\bar{v} + \bar{v}' + v'' + v''') \\ &= \bar{u}\bar{v} + \bar{u}\bar{v}' + \bar{u}v'' + \bar{u}v''' + \bar{u}'\bar{v} + \bar{u}'\bar{v}' + \bar{u}'v'' + \bar{u}'v''' \\ &\quad + u''\bar{v} + u''\bar{v}' + u''v'' + u''v''' + u'''\bar{v} + u'''\bar{v}' + u'''v''' \end{aligned} \tag{3.3.31}$$

and we want to perform an ensemble averaging and vertical integration of this product. Noting that the order in which these are done is arbitrary we first take the ensemble average with the result

$$\begin{aligned} \langle UV \rangle_{\text{ensemble}} &= \bar{u}\bar{v} + \bar{u}v'' + \bar{u}'\bar{v}' + \bar{u}'v''' + u''\bar{v} + u''v'' \\ &\quad + u'''\bar{v}' + u'''v''' \end{aligned} \tag{3.3.32}$$

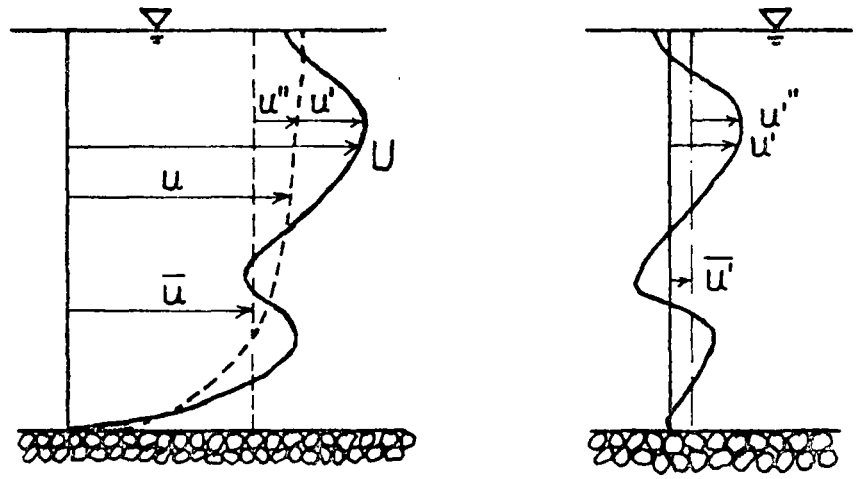


Figure 3-6. Sketch of velocity components.
U = instantaneous local value.
u = ensemble average local value.
$\bar{u}$ = ensemble average, vertical average.
u' = turbulent fluctuation
u" = vertical deviation of u from $\bar{u}$.
$\bar{u}'$ = vertical average of u' .
u'" = vertical deviation of u' from $\bar{u}'$.

since all terms containing only one turbulent fluctuation (') average out to zero. Similarly, an average over depth is carried out.

$$(3.3.33) \qquad \overline{\langle UV \rangle}_{\text{ensemble}} = \bar{u}\bar{v} + \overline{\bar{u}'\bar{v}'} + \overline{u''v''} + \overline{u'''v'''}$$

where overbar means vertical average according to

$$(3.3.34) \qquad \bar{x} = \frac{1}{h+\eta} \int_{-h}^{\eta} x \, dz$$

Again, all terms containing only one vertical deviation (") average to zero. We can now write the total contribution from the convective terms†

$$(3.3.35) \qquad \int_{-h}^{\eta} \langle UV \rangle_{\text{ensemble}} \, dz = \frac{q_x q_y}{H} + \int_{-h}^{\eta} (\langle u'v' + u'''v''' \rangle_{\text{ensemble}} + u''v'') dz$$

The first term in the integral on the right is the usual turbulent Reynolds stress and the two remaining terms are momentum transfers due to the vertical velocity distribution. The integral on the right has so far not been related to the mean flow in a consistent and satisfactory way. Consequently it is often neglected completely. The structure of the terms is similar to the molecular momentum transfer process. But while the latter is a homogeneous isotropic process characterized by the molecular viscosity, this is not the case with turbulent motion and vertical velocity shear. Prandtl used mixing length theory to derive a virtual viscosity for turbulent boundary layer flow [62]. In order to get a closed formulation we postulate a similar functional relationship without

---

† Note that this contribution as in (3.3.15) - (3.3.16) is not strictly correct, because we started out with the ensemble averaged equations.

invoking any mixing length theories.

$$(3.3.36) \qquad F_{x_i x_j} = \int_{-h}^{\eta} \{\tau^{\nu}_{x_i x_j} / \rho_o - \langle (u_i' u_j') + (u_i''' u_j''') \rangle_{\text{ensemble}} - u_i'' u_j''\} \, dz$$

$$= E_{ij} \left(\frac{\partial q_j}{\partial x_i} + \frac{\partial q_i}{\partial x_j} \right) \quad i,j = 1,2 \text{ no summing over } i,j$$

E_{ij} is a symmetric "eddy viscosity" coefficient matrix that depends on the mean flow, depth, applied surface stresses and flow history. What values actually should be used must be determined from experience or by trial since the explicit dependence on the mentioned parameters is unknown. In the literature $1 - 10^{+5} m^2/sec$ have been quoted for the principal values of E_{ij}. In model applications to Massachusetts Bay, the use of values up to 10^4 m^2/sec has apparently not changed the results significantly. In spite of the nebulous circumstances we feel that the inclusion of $F_{x_i x_j}$ has several attractive properties. It allows for internal friction and thereby energy dissipation, provided E_{ij} is positive; it does represent actual physical processes(although not accurately) and it is particularly suitable for damping short wave noise generated by numerical methods.

As an attempt to bring some consistency into the anisotropic case the direction of the local mean current is chosen as the major principal axis of E_{ij} with the minor principal axis perpendicular thereto. This means that in a local coordinate system with the x-axis in the direction of the current, E_{ij} is diagonal:

$$(3.3.37) \qquad E_{ij} = \begin{bmatrix} E_1 & 0 \\ 0 & E_2 \end{bmatrix}$$

The corresponding E_{ij} in the global coordinate system is then found by simple rotation. If θ is the angle from global to local x-axes, (see Figure 3-7) the rotation is written

$$(3.3.38) \qquad E_{ij} = \underset{\sim}{T}^T \begin{bmatrix} E_1 & 0 \\ 0 & E_2 \end{bmatrix} \underset{\sim}{T}$$

where $\underset{\sim}{T}$ is the transformation matrix

$$(3.3.39) \qquad \underset{\sim}{T} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

and superscript T means transpose. In [54] and [75], the ratio of E_1 to E_2 was found to be in the range 10 - 60 for a tidal coast and a lake. Locally negative values of eddy viscosity have been measured indicating energy being fed to the mean flow by turbulent eddies; however, this happens only under very special conditions. For large areas, the overall effect of the internal stresses is to dissipate energy. [14, 67] give a more detailed discussion of this topic with some examples.

We are finally in a situation where we can present a formulation in closed form. For convenience, all the pertinent equations are given below.

Conservation of mass

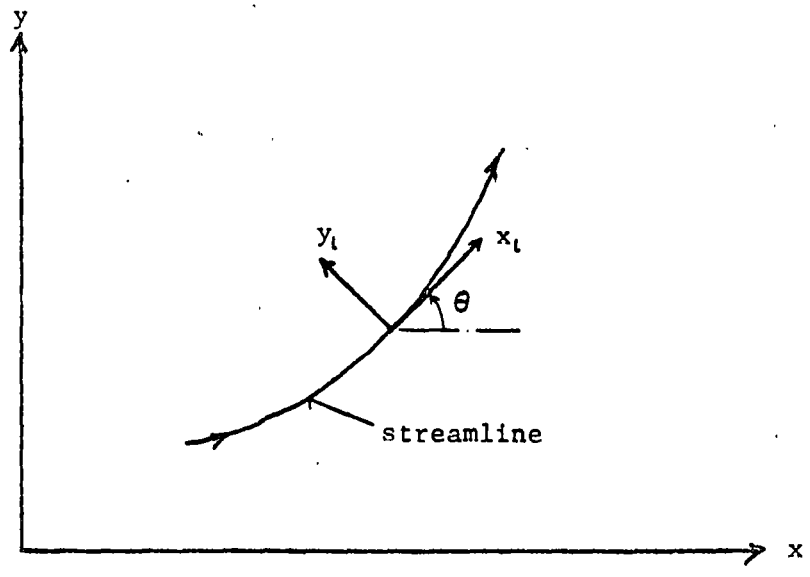


Figure 3-7. Global and local coordinate-system.

(3.3.40) $$H_{,t} + q_{x,x} + q_{y,y} = q_I$$

x and y equilibrium

(3.3.41) $$q_{x,t} + (\bar{u}q_x)_{,x} + (\bar{u}q_y)_{,y} - fq_y + (F_p - F_{xx})_{,x} - F_{yx,y} + \frac{1}{\rho_o}(\tau_x^s - \tau_x^b) - \bar{m}_x - \frac{1}{\rho_o}(p^s H_{,x} + \Delta\rho g H h_{,x}) - g\eta h_{,x} = 0$$

(3.3.42) $$q_{y,t} + (\bar{v}q_x)_{,x} + (\bar{v}q_y)_{,y} + fq_x - F_{xy,x} + (F_p - F_{yy})_{,y} + \frac{1}{\rho_o}(\tau_y^s - \tau_y^b) - \bar{m}_y - \frac{1}{\rho_o}(p^s H_{,y} + \Delta\rho g H h_{,y}) - g\eta h_{,y} = 0$$

with the constitutive relations.

(3.3.14) $$F_p = g h \eta + \frac{1}{2} g \eta^2 + \frac{1}{2} \frac{\Delta\rho}{\rho_o} g H^2 + \frac{p^s}{\rho_o} H$$

(3.3.36) $$F_{x_i x_j} = E_{ij} \left(\frac{\partial q_j}{\partial x_i} + \frac{\partial q_i}{\partial x_j} \right) \quad i,j = 1,2 \text{ no summing over } i,j$$

The bottom and surface shear stresses are given by (3.3.20)-(3.3.21) and (3.3.23) with (3.3.28).

3.4 Boundary Conditions

Defining the correct types of boundary conditions is one of the more critical parts of the formulation process. What prescribed values must be given, and where? The consequence of not specifying enough is normally the existence of non-unique solutions whereas too much may lead to the non-existence of any solution. These issues are often overlooked because the problems are formulated and solved by people who usually do not have

the necessary mathematical background (and time) to worry about the existence and uniqueness of solutions. Still, solutions have been obtained and verified with great success, which probably is due to luck and the fact that generally well behaved physical problems are solved.

In recent years, considerable efforts have been made by mathematicians to prove existence and uniqueness of fluid flow problems, notably solutions of Navier-Stokes equations [34]. Unfortunately, such proofs do not exist yet for our problem and are not likely to be made in the near future. We shall therefore take the "engineering" approach and assume an automatic proof if a reasonable solution is found. To that end, we have to be reasonably certain that the prescribed boundary conditions are proper.

Trying to get a better feeling for what boundary conditions are necessary, we note that the present flow problem is governed by one 2-component vector equation which is the equivalent of Newton's 2nd law:

(3.4.1) Force = mass x acceleration

$$\Downarrow$$

$$F_i = m(x_i)_{,tt}$$

The law of conservation of mass (3.3.4) is thus a constraint to be distinguished from an equilibrium equation.

It is well-known that for a single particle, a solution to (3.4.1) exists and is unique if an initial condition and either the force F_i or the displacement x_i is prescribed. The intuitive generalization to our flow problem is then to specify an initial condition and the force or the discharge which plays the role of displacement in a fluid [4]

at the boundaries. The initial situation is expressed as

(3.4.2) $(q_x, q_y) = (q_{xo}(x,y), q_{yo}(x,y))$ for all (x,y) in Ω and $t = 0$

Ω is the entire interior domain and the initial time is taken as zero. Also the initial mass must be known, thus

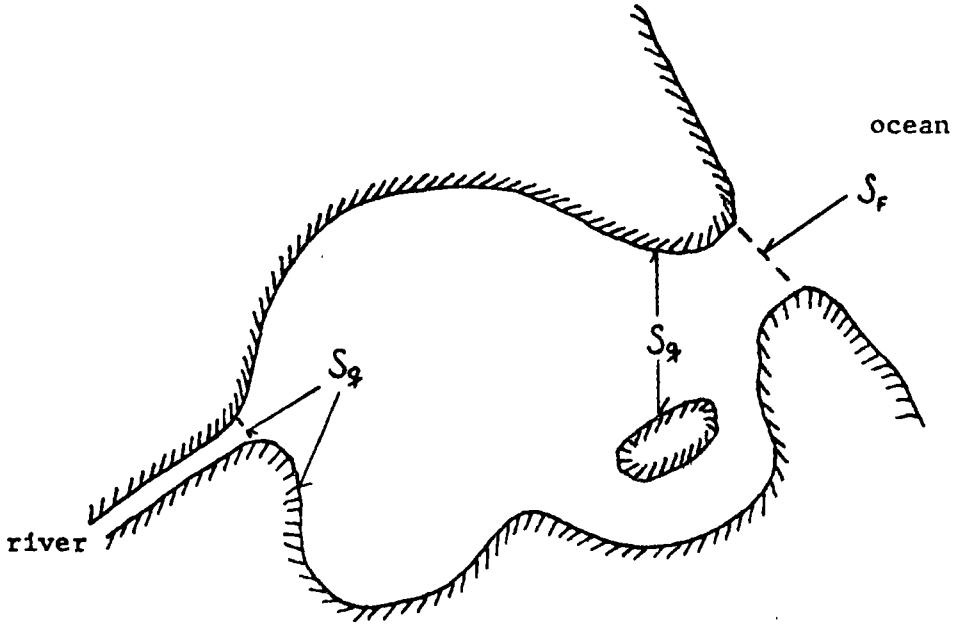


Figure 3-8: Discharge and Force Boundaries

(3.4.3) $H = H_o(x,y)$ for all (x,y) $t = 0$

On the boundaries there are two alternatives as previously mentioned. Referring to Figure 3-8, we distinguish between discharge boundaries S_q and force boundaries S_F. On S_q we write

(3.4.4) $q_n = \alpha_{nx} q_x + \alpha_{ny} q_y = q_n^*$

$$(3.4.5) \qquad q_s = -\alpha_{ny}\, q_x + \alpha_{nx}\, q_y = q_s^* \qquad \text{on } S_q$$

for the normal and tangential discharges, where the direction cosines are

$$(3.4.6) \qquad \alpha_{nx} = \cos(n,x) \; ; \quad \alpha_{ny} = \cos(n,y)$$

and the superscript * signifies a prescribed value.

On the remaining part of the boundary, S_F, the external force, must be given, thus

$$(3.4.7) \qquad F_{nn} = -F_p + \alpha_{nx}^2\, F_{xx} + \alpha_{ny}^2\, F_{yy} + 2\alpha_{nx}\, \alpha_{ny}\, F_{xy} = F_{nn}^* \qquad \text{on } S_F$$

$$(3.4.8) \qquad F_{ns} = (\alpha_{nx}^2 - \alpha_{ny}^2)\, F_{xy} + \alpha_{nx}\, \alpha_{ny}\, (F_{yy} - F_{xx}) = F_{ns}^* \qquad \text{on } S_F$$

must hold for the normal and tangential specific force measures. (Specific force measure is equal to a force per unit width and density).

In the idealized case of an inviscid fluid (3.4.4) and (3.4.7) must still hold, however F_{ns}^* must be zero since no shear can be developed and (3.4.5) can hence not be imposed either.

The continuity equation (3.3.4) is used to find the position of the free surface. It is a mass balance equation and does therefore not require any boundary conditions.

DIRECTORIO DE ASISTENTES AL CURSO USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS(DEL 8 AL 12 DE MAYO DE 1978)

NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
1. ING. MIGUEL ANGEL AGUAYO Y CAMARCO Cultivos No. 144 Col. Progreso del Sur México 13, D. F. Tel: 5-82-61-06	COMISION DE AGUAS DEL VALLE DE MEXICO Balderas No. 55-2o. Piso México 1, D. F. Tel: 5-85-50-66 Ext. 206
2. ING. MOISES BEREZOWSKY VERDUZCO Camino Sta. Teresa 890-XI-304 Col. Contreras México 20, D. F. Tel: 5-68-65-58	INSTITUTO DE INGENIERIA,UNAM Ciudad Universitaria México 20, D. F. México 20, D. F. Tel: 5-50-52-15 Ext. 3608
3. VICENTE BETANZOS VELASCO Norte 69 No. 2923 Col. Popular México 16, D. F.	SECRETARIA DE AGRICULTURA Y Y RECURSOS HIDRAULICOS Plaza de la República No. 31-6 México, D. F. Tel: 5-46-52-75
4. HOMERO R. CABRERA MORA Calle 6a. 123-2 Ensenada, B. C.	CENTRO DE INVESTIGACION Y EDUCACION SUPERIOR DE ENSENADA B. C. Calle Espinoza No. 854 Ensenada, B. C. Tel: 8-13-22
5. ING. OCTAVIO CASTELLANOS LOPEZ Av. Copilco 162 Edif. 2-A-301 Copilco-Universidad México 20, D. F. Tel: 5-50-88-72	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS (CIECCA) Av. San Bernabé 549 San Jerónimo México 20,
6. ING. JORGE LUIS DE VICTORIA ALMEIDA Av. Revolución 820-206 Col. Mixcoac México 19, D. F. Tel: 5-63-90-38	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS (CIECCA) Av. San Bernabé No. 549 San Jerónimo México 20, D. F. Tel: 5-95-24-00

DIRECTORIO DE ASISTENTES AL CURSO USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS (DEL 8 AL 12 DE MAYO DE 1978

NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
7. ING. EUGENIO DOMINGO COBO PEREZ Copilco 300-12-102 Col. Coyoacan México 20, D. F. Tel: 5-50-68-26	INSTITUTO DE INGENIERIA,UNAM Ciudad Universitaria México 20, D. F. Tel: 5-48-97-95
8. JAVIER ESPINOZA CACERES Londres 17 Depto. 203 Coyoacán México 21, D. F.	INSTITUTO DE INGENIERIA,UNAM Ciudad Universitaria México 20, D. F. Tel: 5-50-52-15 Ext. 3607
9. ING. FERNANDO ENSEÑAT MACHADO Viaducto M. Alemán No. 178-9 Col. del Valle México 12, D. F.	DEPARTAMENTO DE ESTUDIOS Y LABORATORIOS- DIRECCION GENERAL DE OBRAS MARITIMAS Lerdo No. 6 San Juan Ixhuatepec Edo. de México Tel: 5-69-28-37
10. LIC. SALVADOR FARRERAS SANZ Apdo. Postal No. 2670 Ensenada, B. C. Tel: 8-13-22	CENTRO DE INVESTIGACION CIENTIFICA Y EDUCACION SUPERIOR DE ENSENADA, B. C. Av. Espinoza 843 Apdo. Postal 2732 Ensenada, B. C. Tel: 8-13-22
11. ING. EDUARDO RAMON FERNANDEZ V. Laguna de la Magdalena No. 430 Col. Ventura Puente Morelia, Mich.	JUNTA DE PLANEACION Y URB. DEL EDO. DE MICH. Casa de Gobierno Libramiento Sur Morelia, Mich. Tel: 2-65-05
12. ING. ARTURO GARCIA MENDOZA Av. Universidad 1810-F-8 Oxtopulco México 20, D. F. Tel: 5-50-01-36	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS (CIECCA) Ave. San Bernabé No. 549 San Jerónimo México 20, D. F. Tel: 5-66-08-88

DIRECTORIO DE ASISTENTES AL CURSO USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS (DEL 8 AL - 12 DE MAYO DE 1978)

NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
13. CARLOS GONZALEZ GUZMAN Dakota 395-5 Col. Nápoles México 18, D. F. Tel: 5-36-62-69	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS (CIECCA) Ave. San Bernabé 549 San Jerónimo México 20, D. F. Tel: 5-95-44-53
14. JESUS MANUEL HAM CHI Libra No. 19 Depto. 4 Col. Prado Churubusco México 13, D. F. Tel: 6-70-18-87	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS Paseo de la Reforma No. 69-4o Piso Col. Juárez México 1, D. F. Tel: 5-35-25-25
15. ING. ENRIQUE C. HERNANDEZ CORTES Lidia 88-1 Col. Guadalupe Tepeyac México 14, D. F. Tel: 5-37-09-58	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS Paseo de la Reforma 107-8o.P. México 1, D. F. Tel: 5-66-06-88-117
16. ING. MANUEL A. HUIDOBRO GARCIA Cartago No. 88 Lomas Estrella México 13, D. F.	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS (CIECCA) Ave. San Bernabé No. 549 San Jerónimo México 20, D. F. Tel: 5-95-44-53
17. ING. HUMBERTO JIMENEZ DIAZ Zacatecas No. 33-5 Col. Roma México 7, D. F. Tel: 5-84-87-33	DEPARTAMENTO DE ESTUDIOS Y LABORATORIOS DIRECCION GENERAL DE OBRAS MARITIMAS Lerdo de Tejada No. 6 Col. Marina Nacional Edo. de México Tel: 5-69-28-36

DIRECTORIO DE ASISTENTES AL CURSO USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS (DEL 8 AL 12 DE MAYO DE 1978)

NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
18. ING. EDUARDO LOZANO GONZALEZ Paseo de la Reforma No. 107-8o. Piso Col. San Rafael México 4, D. F. Tel: 5-46-14-55	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS (CIECCA) Paseo de la Reforma 107-1er. Col. San Rafael México 4, D. F. Tel: 5-66-06-88 Ext. 140
19. ING. GUSTAVO LUNA ESCALANTE San Antonio 134-23 Col. Nápoles México 18, D. F.	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS SUBDIRECCION DE PROMOCION Y PROGRAMAS Paseo de la Reforma No. 35-10 Piso México 1, D. F. Tel: 5-92-33-24
20. JESUS MAGALLANES PATIÑO Ezequiel Montes 120-24 Col. San Rafael México 4, D. F. Tel: 5-46-13-50	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS Paseo de la Reforma 69-4o.Piso México, D. F. Tel: 5-46-95-20
21. FRANCISCO J. MAYTORENA FONTES Bucareli 80 Int. "M" México 1, D. F. Tel: 5-12-68-07	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS Paseo de la Reforma No. 107 México 1, D. F. Tel: 5-92-10-31
22. JESUS R. MENDOZA RUIZ Albino García No. 72 Col. Vista Alegre México, D. F. Tel: 5-19-04-83	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS Plaza de la República No. 31 Col. Tabacalera México 8, D. F. Tel: 5-46-50-96
23. CARLOS ANGEL Q. MORTERA GUTIERREZ Av. Unión 281 Col. Tepeyac Insurgentes México 14, D. F. Tel: 5-77-62-35	DEPARTAMENTO DE ESTUDIOS Y LABORATORIOS-DIRECCION GENERAL DE OBRAS MARITIMAS San Juan Ixhuactepec Tel: 5-69-50-30

DIRECTORIO DE ASISTENTES AL CURSO USO DE COMPUTADORAS EN PROBLEMAS DE CIRCULACION Y DISPERSION EN AGUAS COSTERAS, LAGOS Y RIOS (DEL 8 AL 12 DE MAYO DE 1978)

	NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
24.	ING. ARMANDO MUÑOZ PARGA Laures No. 9 Col. Sta. Ma. Ribera México, D.F. Tel. 547-32-48	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS Plaza de la República No. 31 6o. Piso México, D.F. Tel. 549-50-96
25.	ING. VICTOR S. PINEDA ESPINOSA Río Becerra No. 473-101 Col. Nápoles México 18, D.F. Tel. 543-82-83	DEPARTAMENTO DE ESTUDIOS Y LABORATORIOS-DIR. GRAL. DE OBRAS MARITIMAS Lerdo de Tejada No. 6 Col. Marina Nacional Edo. de México Tel. 569-28-36
26.	EUGENIO RIQUELME TORRENTE Ret. 10 Dr. N. León Gpo. 18-N Col. J. Balbuena México 9, D.F. Tel. 552-40	SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS (CIECCA) San Bernabé No. 549 San Jerónimo México 20, D.F. Tel. 595-53-44
27.	ING. HONORIO RIVERA MOCTEZUMA Lago Tana No. 66-C Torre Blanca México 17, D.F.	CENTRO DE EDUCACION CONTINUA Tacuba No. 5-1er. Piso México 1, D.F. Tel. 521-40-20
28.	ING. FRANCISCO ROMERO LUNA Av. Cuauhtémoc No. 883-10 Col. Narvarte México 12, D.F. Tel. 543-63-60	INSTITUTO DE INGENIERIA, UNAM Ciudad Universitaria México 20, D.F.
29.	RAFAEL F. SAENGER Y FERNANDEZ López Cotilla No. 756 Col. Del Valle México 12, D.F. Tel. 523-52-53	INSTITUTO DE INGENIERIA, UNAM Ciudad Universitaria México 20, D.F. Tel. 550-52-15 Ext. 3610
30.	ING. ANTONIO YOKOYAMA KANO Torquemada No. 42 Col. Obrera México 8, D.F.	UNIVERSIDAD CATOLICA MADRE Y MAESTRA Autopista Duarte Santiago, República Dominicana