Forecasting Research

Met O 11 Technical Note No. 11

"Panel-Beater" : A proposed fast algorithm for semi-geostrophic finite-element codes

by

R. J. Purser

June 1988

Me O 11)
London Road, G12 2SZ, England

JM

MET O 11 TECHNICAL NOTE NO 11

"PANEL-BEATER": A PROPOSED FAST ALGORITHM FOR SEMI-GEOSTROPHIC FINITE-ELEMENT CODES

by

R J Purser

LONDON, METEOROLOGICAL OFFICE.
Met.O.11 Technical Note (New Series) No.11

"Panel-beater": a proposed fast algorithm for semi-geostrophic finite-element codes.

02320888 FH2A

Met O 11 (Forecasting Research)
Meteorological Office
London Road
Bracknell
Berkshire RG12 2SZ
England

June 1988

ABSTRACT

Integration of the semi-geostrophic equations in their conservative finite elements form has hitherto been hindered by the lack of an efficient algorithm to analyse the topology and metrical parameters of each instantaneous solution. With existing techniques it has been found that in large problems the computation of the geometrical configuration of the elements consumes an overwhelming proportion of the total CPU time, putting the investigation of three dimensional problems by the finite element method beyond the reach of even the fastest computers for problems of sufficient size to be interesting. A new and very different algorithm is proposed here which is expected to be able to accomplish the geometrical constructions in a reasonable time for even quite large problems. The new method, which will be referred to as the "Panel-beater" (PB) algorithm, exploits an adaptive or evolutionary approach rather than a purely constructive one, relying on an internally consistent previous construction which is transformed through a sequence of local mutations towards the true topological configuration. For time-dependent problems the starting point of the algorithm at each time level is the final construction valid at the previous time level. In this way the relatively minor topological modifications one might expect during a single time interval should be reflected in a typically modest computational expenditure on this part of the code.

1. INTRODUCTION

Following the proposal by Cullen (1983) to use a finite element technique in the computation of semi-geostrophic (SG) problems, much progress has been made in this area, notably in the case of two-dimensional solutions computed by Chynoweth (1987). In a geometrical analysis of the f-plane SG theory Cullen and Purser (1984) (henceforth denoted (CP) established that each dynamically stable solution, in either two or three dimensions, could be uniquely associated with a certain convex function of the space coordinates in such a way that, by introducing an auxiliary coordinate measuring the value of this function a generalized curvature of the manifold constituting the solution in the extended space could be immediately identified with the SG potential vorticity of the standard theory (Hoskins and Bretherton, 1972; Hoskins 1975). It was further recognised that finite element or "geometric" solutions would then take the form of polyhedral surfaces, convex from the direction of negative auxiliary coordinate, in the extended space and, it was argued, that the subsequent evolution of such a solution would be dictated only by each element's centroid values of geostrophic wind and by boundary constraints that together could be expressed as a finite coupled set of ordinary differential equations.

The geometrical method has been applied in a variety of quasi-static two-dimensional solutions (Cullen et al., 1987a; Cullen et al., 1987b) and extended to axi-symmetric solutions by Shutts et al. (1988). However, attempts to implement the geometric method for large problems, even in only two dimensions, have demonstrated that the existing algorithms for constructing the geometry of each solution consume an increasingly disproportionate amount of the total CPU time as the size of the problem grows (Chynoweth, 1987). Therefore it appears that the present method is impractical for investigating the necessarily large geometrical constructions required to represent interesting three-dimensional flows such as those associated with evolving large amplitude baroclinic waves.

In any finite element construction for SG theory from "data" comprising a list of the size, potential temperature and geostrophic momentum values of each fluid element, the first step is always a tabulation of the vertices of the solution and their connecting edges. Fortunately one finds that for typical data possessing no contrived symmetry the fluid elements or "cells" meet at vertices three at a time in two-dimensional problems, four at a time in three-dimensions, except for the fleeting moments at which the connectivity of vertices and edges that delineate the solution switch as edges vanish or emerge at points. The exceptional cases containing these singular events can be safely ignored in practice (since such topologies are not stable to infinitesimal perturbations of the data their appearance numerically in code that uses 64 bit arithmetic would require an exceptional coincidence). Thus it is natural to express the connectivity of the polyhedral solution at each instant by creating a table of vertices, each identified uniquely by its three or four adjacent numbered fluid or boundary cells (listed in increasing order) together with entries to point to the three or four neighbouring vertices one arrives at by following the edges radiating from this vertex in a direction opposite to each adjacent cell (stored in corresponding order). A segment of such a tabulation is shown in Figure 1a together with a schematic depiction of the corresponding portion of the geometry (Figure 1b) for a two-dimensional construction. The excessive computation required to create such a tabulation in problems with a large number, say N, elements is attributed to the fact that to determine the identity and coordinates of the vertex terminating a newly found edge involves an exhaustive search of the feasible intersections with all other element-hyperplanes implied by the data until the nearest intersection in the positive direction is found — that being the actual one sought. Assuming the number of edges is roughly proportional to N and each requires the computation of almost N intersections in order to ascertain the closest, this basic step of the code will consume order N^2 computations. Elsewhere in the code computational expenditure appears not to exceed order N log N so the principal bottleneck would indeed appear to be in the construction of solution-vertices for large problems.

Even if a given solution needs to be perturbed only slightly — typically providing a new network of edges and vertices that differs topologically from the old in only a minority of the connections, the algorithm as discussed above is unable to exploit persisting similarities to short-cut the new construction. In an attempt to avoid the bottleneck discussed above an alternative approach, which does use prior information about likely connections, is proposed. The geometrical details will be given in the next section and it will suffice here to state that the essence of the new method is to adapt the geometry by a succession of local reconnections which are interpreted in the Legendre-dual representation of the problem (Purser and Cullen 1987) as the successive removal of local "dents" or concavities through reconstructions that increase the size of the extended space "above" (in the sense of increasing auxiliary coordinate, R) the dual solution surface R(X) (i.e., enlarging the "epigraph" of R). It is only in the dual representation that the name "panel-beater" is appropriate as in the physical space the process is rather one of undoing a tangle of self-intersections of the solution surface. The most significant feature of the new method is that, if it proceeds successfully, the effort involved in performing the transformation from one feasible solution to the actual solution should in typical circumstances involve order N log N, not N^2, computations, thereby circumventing the bottleneck that has so far rendered the solution of large problems unattainable.

2. Stable properties of finite-element solutions of semi-geostrophic theory

The notation is that of Purser and Cullen (1987), henceforth denoted PC, who also discuss at length the physical interpretation of the geometric duality we refer to. The "data" are defined by N values of dual coordinates $\underset{\sim}{X}_\alpha$ and N values of mass M_α together with a physical domain D with exactly the area or volume M_α that contains the given fluid elements.

$$M_D = \sum_{\alpha=1}^{N} M_\alpha \qquad (2.1)$$

The "solution" consists of a continuous function $P(\underset{\sim}{x})$ defined for each

point $\underset{\sim}{x}$ in D which can be divided onto the N fluid elements or "cells" C_α such that within each cell,

$$\underset{\sim}{X}_\alpha = \nabla_{\underset{\sim}{x}} P \quad \underset{\sim}{x} \quad \varepsilon \; C_\alpha \quad , \tag{2.2}$$

and the size of the cell is as given by the data, i.e., in three dimensions:

$$M_\alpha = \int_{C_\alpha} dx\, dy\, dz \equiv \int_{C_\alpha} d\mu \; . \tag{2.3}$$

CP showed that the solution for given data and convex domain D is unique up to an additive constant for P and that the elements C_α are each convex polygons or polyhedra.

Recall that a convex set is one that wholly contains any line segments whose end points lie in the set. A convex function $P(\underset{\sim}{x})$ defines a convex set in the extended space $(\underset{\sim}{x}, s)$ by its epigraph,

$$\text{Epi}(P) \equiv \{(\underset{\sim}{x}, s): \; s \geq P(\underset{\sim}{x})\} \quad . \tag{2.4}$$

PC noted the standard result from convexity theory (e.g., Rockafellar 1970) that the "dual" function

$$R(\underset{\sim}{X}(\underset{\sim}{x})) = \underset{\sim}{x} . \underset{\sim}{X} - P(\underset{\sim}{x}) \tag{2.5}$$

is a convex function of $\underset{\sim}{X}$ if P is a convex function of $\underset{\sim}{x}$. Of interest here is the fact that if the solution P is made up of polygons or polyhedra, the dual function R is also a "polyhedral" function, but the cells of R correspond to the vertices of P and vice-versa as illustrated in Figure 2. Note from the definition of R that it is the negative of the P coordinate at $\underset{\sim}{x} = 0$ of the hyperplane tangent to the solution P at $\underset{\sim}{x}$.

The construction of the polyhedral configuration of P resulting from specified values of $\underset{\sim}{X}_\alpha$ and R_α is all that concerns us in this note — the derivation of the R_α from the masses M_α is a separate problem discussed in CP and Chynoweth (1987) and involves the inversion of a large non-linear coupled system, but one which can be satisfactorily handled by standard numerical techniques. Note that the duality associates with every hyperplane of the extended physical space $(\underset{\sim}{x},P)$ a point of the extended dual space $(\underset{\sim}{X},R)$ and vice-versa, including points "at infinity" along given along given rays if hyperplanes are parallel to the P-axis. In this way the rigid walls of the boundary of D can be mapped to

points at infinity in ($\underset{\sim}{X}$, R)-space. With the further modification of the geometrical solution obtained by imposing a "lid" consisting of an inverted rectangular or cuboidal element to close off the polytope in ($\underset{\sim}{x}$, P) space at a suitably large value of P, as shown schematically in cross-section in Figure 3, we may assert that for typical data containing no intrinsic symmetry all vertices of the solution are associated with exactly n+1 edges, where n is the dimensionality of the physical domain. To formalise and justify this assertion requires some further definitions.

Definition 1: An "n-simplex" is generically the shape of the convex hull of n+1 points not lying on the same hyperplane in R^n. It is exemplified in the cases n=2 and n=3 by triangles and tetrahedra respectively.

Definition 2: A "graph" in general is a logical structure comprising a set V of "nodes" v_i together with a set E of "pointers" e_j, each comprising an ordered pair of nodes. Such a pointer, e.g., (v_i, v_j) is said to "point" from v_i to v_j. Here we shall restrict ourselves to pointer connections that are mutual, i.e., (v_i, v_j) implies the existence of (v_j, v_i), and also we assume no node points to itself.

Each geometric solution is associated with a particular graph by representing vertices of the polyhedral surface ($\underset{\sim}{x}$, P), including the lid, by nodes and representing edges by mutual pointers. Note that the graph does not contain metrical information — thus the vertex-nodes v_i might be identified solely by the set of adjacent fluid, boundary or lid elements that meet there — but not by coordinates. The metrical details are determined by what we shall refer to as the "configuration".

Definition 3: A "configuration" $\underset{\sim}{\chi}$ is defined to be the vector in q-dimensional "configuration space" $S \equiv R^q$ with components formed from the values $\underset{\sim}{X}_\alpha$ and R_α at all the fluid elements α. In an n-dimensional domain with $\underset{\sim}{N}$ fluid elements q=N*(n+1). Each configuration $\underset{\sim}{\chi}$ is implicitly associated with a particular graph $G(\underset{\sim}{\chi})$ formed from the identities of vertices and edges of the convex geometric solution corresponding to $\underset{\sim}{X}$.

Definition 4: A configuration is defined to be "stable" or "non-degenerate" if its graph remains the same for all variations of χ within a sufficiently small neighbourhood in S.
We may now formalize our previous assertions:

Theorem 1: For solutions in $\underset{\sim}{x} \in R^n$ the only stable configurations are those for which each vertex v_i is connected to precisely n+1 others.

Theorem 2: The set of configurations χ corresponding to non-stable graphs has measure zero in S.

Remarks: Theorem 1 identifies stable configurations with solutions whose elements meet three or four at a time according to whether the problem is set in two or three dimensions respectively. By implication, the dual solutions of stable configurations comprise polyhedral surfaces whose cells are each an n-simplex.

Theorem 2 states that non-stable configurations are "rare" in the sense that a configurations chosen at random from S will be stable almost always.

Proof: Suppose a cell of dual solution R is not a simplex. Then one of its vertices, not a point at infinity, may be perturbed by an arbitrarily small amount in the direction of negative R to displace this point off the plane determined by the other vertices, thereby requiring a change in connectivity to re-establish a convex surface. If instead the connectivity is preserved by manipulating the R or $\underset{\sim}{X}$ of other vertices of the cell then at least one linear constraint among the $\underset{\sim}{X}$ and R of the vertices must be imposed, implying the set of non-stable configurations can at most occupy a countable set of "flat" patches of dimensionility less than q.

Definition 5: A non-stable configuration χ is defined to be "rth-order degenerate" if the set S' of configurations with the same graph as $G(\chi)$ defines a patch of dimensionality q-r in S. We will denote by S_r the set of rth order degenerate configurations.

Theorem 3: Every first-order-degenerate configuration contains precisely one vertex of the solution P (or cell of solution R) that involves n+2 connections, all other vertices (or cells of solution R) being of the standard (n+1)-type.

Proof: A configuration implying a single (n+2)-connection vertex requires a single linear constraint among the $\underset{\sim}{X}$ and R components to preserve the co-planarity of the vertices of the non-standard cell of dual space. If more than n+2 vertices of a cell were required to be coplanar in dual space, or if more than one cell were required to preserve n+2 vertices coplanar, one linear constraint in configuration space would not suffice.

Evolutionary aspects of solutions are most conveniently studied by considering each evolution as a "world-line" or continuous curve in the augmented space $S^* \equiv S \times R^1$ obtained by including the time dimension. Suppose world-line W is a member of a family of such evolutions with $\tau_1, \ldots \tau_p \equiv \underset{\sim}{\tau}$ parametrizing the family smoothly in the sense that for suitable vector norms of the gradient operator $\frac{\partial}{\partial \tau}$:

$$\left|\left| \frac{\partial}{\partial \tau_i} \chi(t) \right|\right| < \infty,$$

$$\left|\left| \frac{\partial}{\partial \tau_i} \left(\frac{d\chi}{dt}\right) \right|\right| < \infty, \qquad \chi(t) \in W(\underset{\sim}{\tau}).$$

The evolution $W(\underset{\sim}{\tau}_o)$ is defined to be "stable" at $\underset{\sim}{\tau}_o$ if for all such parametrizations there exists a neighbourhood in $\underset{\sim}{\tau}$ space about $\underset{\sim}{\tau}_o$ sufficiently small that all evolutions corresponding to parameters $\underset{\sim}{\tau}_i$ within this neighbourhood give rise to the same sequence of graphs.

Theorem 4: If an evolution is stable, all the degenerate configurations encountered are first-order and of vanishing duration.

Theorem 5: If the evolution $W(\tau_o)$ is non-stable or "degenerate" there exists a smooth parametrization (in the sense defined above) of neighbouring evolutions in which the set of $\underset{\sim}{\tau}$ corresponding to the same sequence of graphs possesses zero measure in $\underset{\sim}{\tau}$-space.

Remarks: the consequences of theorems 4 and 5 as they relate to stable evolutions are analogous to the consequences of theorems 1 and 2 as they relate to stable configurations. Thus an evolution of a geometric solution in which a "transition" of the connectivity between one stable form and another is not by way of a first order degenerate configuration or in which a transition is not of vanishing duration is regarded as inconceivably rare.

Proof: We rely on a geometrical result discussed by Gilmore (1981) concerning the "transversality" of intersections of submainfolds. A curve $W(\underset{\sim}{\tau}_o)$ intersects the patches constituting the prism-set $S^*_1=(S_1,t)$ "transversally" if it penetrates each patch non-tangentially. Smooth perturbations of W then do not change the sequences of graphs implied because the dimensionality of each patch (= q) and the dimensionality of the curve (=1) sum to the dimensionality of S* (= q+1), i.e., an intersection is still obtained with each patch regardless of the direction of the local perturbation in $\underset{\sim}{\tau}$-space. On the other hand, if W belongs to a degenerate family of evolutions at least one intersection of the curve W must be within a patch in S_r*, r>1 of lower dimensionality than q. There are then infinitesimal, but smooth, perturbations of W that will cause it to "miss" this patch, implying a different sequence of graphs. (N.B. an example of a non-transversal intersection is that of two smooth curves intersecting non-tangentially in three dimensions since a small perturbation of either curve can cause them to miss each other. In the case of a curve intersecting a surface non-tangentially in three dimensions a small perturbation will merely move the intersection, but will not cause it to disappear.)

The results thus far have been obtained for configurations defined as sets of given $(R_\alpha, \underset{\sim}{X}_\alpha)$. In order to show that the same results hold for "data" defined as given sets of $(M_\alpha,\underset{\sim}{X}_\alpha)$ it is necessary to show that the transformation from $s_\alpha = -R_\alpha$ to M_β at fixed $\underset{\sim}{X}_\gamma$ is effectively unique (as

proved in CP) but also non-degenerate. When these conditions are satisfied the property of transversality of intersections is preserved. To demonstrate the non-degeneracy of the mapping we first prove:

Lemma: Suppose a real symmetric matrix J possesses the following properties of diagonal dominance and connectedness,

i) For all α,

$$J_{\alpha\alpha} \geq \sum_{\beta \neq \alpha} | J_{\alpha\beta} | \quad . \tag{2.6}$$

ii) There exists at least one α for which

$$J_{\alpha\alpha} > \sum_{\beta \neq \alpha} | J_{\alpha\beta} | \tag{2.7}$$

iii) Every distinct pair (α,β) may be linked by a "path",

$$\alpha_o, \alpha_1 \cdots \alpha_p$$

with $\alpha_o \equiv \alpha$, $\alpha_p \equiv \beta$ such that

$$J_{\alpha_i, \alpha_{i+1}} \neq 0 \quad i = 0, 1 \cdots p-1. \tag{2.8}$$

Then provided condition (i), (ii) and (iii) are satisfied J is positive definite (all eigenvalues strictly positive)

Proof: Let V be any eigenvector of J, λ its eigenvalue, and partition the set of component indices,

$$A \equiv \{\alpha: \ |V_\alpha| = \max_\beta V_\beta\} \tag{2.9a}$$

$$B \equiv \{\alpha: \ |V_\alpha| < \max_\beta V_\beta\} \tag{2.9b}$$

If B is empty then by property (ii) there is an $\alpha \in A$ such that, normalising $V_\alpha > 0$,

$$\sum_{\beta \neq \alpha} | J_{\alpha\beta} V_\beta | \leq \sum_{\beta \neq \alpha} |J_{\alpha\beta}| V_\alpha < J_{\alpha\alpha} V_\alpha \qquad (2.10)$$

Hence,

$$\sum_{\beta} J_{\alpha\beta} V_\beta = \lambda V_\alpha > 0 \qquad (2.11).$$

If B is not empty then by property (iii) there is an $\alpha \in A$ and a $\beta \in B$ such that $J_{\alpha\beta} \neq 0$ and hence , with normalisation $V_\alpha > 0$,

$$\lambda V_\alpha = \sum_{\beta \in A} J_{\alpha\beta} V_\beta + \sum_{\beta \in B} J_{\alpha\beta} V_\beta > \sum_{\beta \in A} J_{\alpha\beta} V_\alpha - \sum_{\beta \in B} |J_{\alpha\beta}| V_\alpha \geq 0 \qquad (2.12)$$

Thus in either case, $\lambda > 0$ and the lemma is proved. –

Theorem 6: For finite element SG problems in a convex domain with positive measure M allocated to each element and distinct gradients $\underset{\sim}{X}$ the vector of measures $\underset{\sim}{M}$ is a unique differentiable function of the vector of intercepts $s_\alpha \equiv - R_\alpha$ provided a single component of the latter, say s_N is fixed. Moreover, the Jacobian matrix of differentials is non-singular so that solutions $\underset{\sim}{s}$ are implicitly functions of the data $\underset{\sim}{M}$.

Proof: Suppose n=3. As shown in CP the sensitivity of volume M_α to a change in intercept of neighbouring element β is

$$\frac{\partial M_\alpha}{\partial s_\beta} = - \frac{a_{\alpha\beta}}{|\underset{\sim}{X}_\alpha - \underset{\sim}{X}_\beta|} = \frac{\partial M_\beta}{\partial s_\alpha} \quad , \qquad (2.13)$$

where $a_{\alpha\beta}$ is the area of the interface. Diagonal terms are given by

$$\frac{\partial \underline{M_\alpha}}{\partial s_\alpha} = - \sum_{\beta \neq \alpha} \frac{\partial \underline{M_\alpha}}{\partial s_\beta} \quad , \qquad (2.14)$$

where summation includes element N. Choose the constrained element N to be one that does not divide the remaining domain into disconnected portions. Then the matrix J of order N-1 formed from the differentials above by ignoring row and column N satisfy each of the properties of the lemma: property (i) follows immediately from (2.14), property (ii) from the fact that element N makes finite contact with at least one other element, property (iii) from the connectedness of the region of domain D excluding element N. The corresponding proof for two dimensional problems follows trivially.

To summarise the significance of the results obtained in this section, we have shown that instantaneous finite element configurations are almost always characterised by n+1 elements touching at any vertex. Any exceptions are atypical and possess tabulations of the connectivity (i.e. graphs) which are not stable to infinitesimal perturbations of the data. However, when considering a complete evolution there will be numerous instants when the connectivity undergoes a transition through unstable states that are of the non-standard kind, but then these transitional states themselves are almost certainly of the first-order kind involving precisely one non-standard vertex where n+2 elements touch. It is therefore fair to argue that all evolving finite element solutions of interest can be regarded as fitting the standard pattern of belonging to one particular connectivity for a period before instantaneously undergoing a first-order transition to a new connectivity, and so on. The regular "snapshots" constituting the successive states of a discretized model can then be assumed to be of standard form. In principle, if the timesteps were reduced sufficiently one could infer that the change in connectivity from one step to the next could be explained in each case by a simple first-order transition but in practice it would be wise to base any algorithm on the assumption that several transitions have occurred during any given timestep. The following section will consider the means by which the old connectivity may be recognised to be in error and, if necessary, adapted by a succession of first-order transitions to the appropriate connectivity for the configuration.

3. Symbolic panel beating

a. The normal case

A cell of a standard dual solution in n-dimensions can be denoted by the n+1 identifying fluid or boundary elements constituting that cell's vertices, listed in lexicographic order for easy recognition as illustrated in Fig 1a. For example, in a two-dimensional solution two connected vertices of the physical solution, and hence two adjacent triangular cells of the dual solution might be denoted ABC and ABD, AB being the joining edge in physical space or the separating interface in dual space, as depicted in Figures 4a and 4b respectively. The surface R formed by the cells is two-sided, allowing each cell to be assigned an orientation. For the cells in the shape of a simplex this orientation is conveniently signified by chiral prefix "+" or "-" according to whether the order of the vertices given is right handed or left handed when the dual coordinates (X, Y) or (X, Y, Z) are a right-handed set. Thus the oriented triangles of Figure 4a would be denoted +ABC, -ABD. With these definitions the proper chirality may be determined from the sign of the determinant of the matrix of columns formed by appending the $\underset{\sim}{X}$-coordinates of successive vertices (in order) to first row values of unity. For example, in two-dimensions

$$U_{ABC} = \begin{bmatrix} 1 & 1 & 1 \\ \underset{\sim}{X}_A & \underset{\sim}{X}_B & \underset{\sim}{X}_C \end{bmatrix} , \qquad (3.1)$$

and the sign of $|+ABC| \equiv \det (U_{ABC})$ provides the appropriate prefix. Note that this determinant is proportional to the projected area of the triangle ABC. By a standard property of determinants the chiral signature of a given oriented simplex is reversed if, relaxing the lexicographic ordering convention, we transpose a pair of vertices. In general odd permutations reverse the chiral signature while even permutations preserve it. Thus the following six ways of representing the oriented triangle are equivalent:

+ABC, +BCA, +CAB,
-CBA, -ACB, -BAC. (3.2)

Similarly, an oriented tetrahedron can be presented in 24 different but equivalent ways.

We now consider the symbolic manipulations that represent first order transitions of the connectivity of solutions. Suppose the $\underset{\sim}{x}$-space edge between vertices +ABC, -ABD of the physical solution displayed schematically in Fig 4a evolves through the transition (Figure 4c) to a new connectivity with vertices -ACD, +BCD linked by an edge (Figure 4e). The states in the dual representations corresponding to those of Figure 4 a, c, e are respectively those of Figure 4 b, d, f. If we denote by +ABCD* the oriented tetrahedron in the extended space of coordinates (X,Y,R) then, because of the necessity for surface R to be convex in (X,Y), the signed volume of this tetrahedron changes from positive before transition to negative after transition. This latter volume is proportional to the determinant generalising (3.1) which we write as |+ABCD*| , the asterisk denoting the involvement of the R-dimension. This pattern may be made the basis for the simplest application of the panel-beater algorithm in two dimensional problems as codified in the following basic steps:

i) Examine the next of dual-space triangles that were adjacent at the preceding iteration.

ii) Arrange the vertices of the dual space triangles (temporarily) into the pattern +ABC, -ABD by means of signed permutations of the form (3.2)

iii) Determine the sign of |+ABCD*| and return to (i) if positive but

iv) if negative, replace +ABC, -ABD by new triangles equivalent to -ACD, +BCD (making all necessary "housekeeping" alterations to the various tables needed to record the connectivity). Then return to step (i) to consider the next untested pair of triangles.

It may occur that, having made the transition it is found that one of the new triangles, say -ACD occupied the space of its chiral opposite, i.e. +ACD, which already exists, thereby forming what we shall refer to as a "pocket". In that case the appropriate action is to annihilate both members of the pair together with dual space vertex A itself. Physically this might correspond to the transition in which triangular fluid element A has shrunk to a point and vanished, being replaced by the single remaining vertex, +BCD in this case, as shown in Fig 5. This pair-annihilation rule should therefore be appended to step (iv) above.

The generalization of this basic panel-beater algorithm to three dimensions is fairly straightforward but now the pair-annihilation process need not be accompanied by the loss of a dual-space vertex. Also, without annihilation the removal of two tetrahedra is accompanied by the creation of three new ones, so clearly pair annihilation in the n=3 algorithm is a natural result of approximately half the transitions of a typical evolution. The steps of this algorithm are written out as follows:

(i) Examine the next of dual-space tetrahedra adjacent at the preceding iteration.

(ii) Arrange the vertices of each pair of dual-space tetrahedra that were adjacent at the preceding iteration into the pattern +ABCD, -ABCE by means of signed permutations.

(iii) Determine the sign of $|+ABCDE^*|$ and return to (i) if positive, but

(iv) if negative, replace +ABCD, -ABCE by new tetrahedra -ABDE, +ACDE,-BCDE and, if necessary, carry out any pair annihilations necessary if one or more of +ABDE, -ACDE, +BCDE exist already. Then return to (i).

In the case of both two and three dimensional transitions, the action of step (iv) in each case is to detect a location on the surface R that is locally "hollow" and to "fill" it, by adding an (n+1)-simplex to epi(R), thereby increasing the size of the extended ($\underset{\sim}{X}$,R) space occupied by epi(R). Figure 6 illustrates the transition from the two tetrahedra to three implied by step (iv) in "exploded" views of the cells involved.

b. Snags and their resolution

It is theoretically possible for the basic panel-beater algorithms described in (a) to converge to a nonsensical solution if, within the preceding time-step the deformation has been such as to cause a relatively major rearrangement of dual-vertices, implying that actually several transitions of the graph have occurred. In such cases it is possible for the algorithm to converge erroneously to a state in which the oriented surface R has curled in on itself and become self-intersecting even though the tests of step (iii) applied to each dual-cell interface detect no concavity. An illustration of the relevant portion of a state of this kind, with its "shadow", is given in Figure 7a for triangles (n=2). Note that all "dihedral angles" between adjacent elements are positive as registered by step (iii). However, at least one of the triangular elements itself is improperly oriented with respect to increasing R (i.e, inverted). It is therefore necessary to check for proper orientation of each simplex of the dual solution in order to verify that a self-intersecting or "snagged" solution has not inadvertently been created.

We note that if there is any inverted element in the dual solution, then the solution function R will be at least three-valued at that location of X-space. Also, it will be found that at least one of the inverted element's vertices will correspond to a location of $\underset{\sim}{X}$-space where the function is at least double valued. The intention is to reduce the solution by iterative steps to one which is single valued in R by first removing such a vertex and its neighbouring elements and replacing the gap by a covering of new triangles or tetrahedra. The edges (interfaces) of

this new covering can then be checked for positive dihedral and adjusted if necessary, while the value R of the vertex which was removed can be compared with the value of the properly oriented element's R at the same location. If the vertex has a lower value of R, then it is re-inserted inside the comparison element as the common vertex of three new elements there. Note that the search for the comparison element is not necessarily confined to the immediate (topological) neighbourhood of the original vertex, so this process could be computationally costly.

For brevity we shall drop chiral prefixes and lexicographic ordering here and write elements of the R-surface using a right-handed ordering of their vertices relative to the orientation of this surface. To illustrate the procedure consider the portion of the two-dimensional solution illustrated in Figure 7a which contains two inverted elements, BED, CDE and two vertices, D and E, corresponding to positions $\underset{\sim}{X}$ where the surface is double valued in R. The projected "shadow" of the configuration is shown in the lower portion of the figure and the six triangular facets comprising this somewhat convoluted surface are listed (in right-handed notation) at the left of the figure. We shall illustrate the resolution of this snagged surface by considering two possible strategies.

First, suppose we consider the consequence of removing the offending vertex D and replace the three adjoining elements by the single new element BEC. Since R(D) > R(D'), where D' is the projection of D on element BFE, we do not reinstate D as a vertex of the solution. This stage is illustrated in Figure 7b. However, we have not finished yet as the new edge BC between ABC and BEC has negative dihedral and so using the standard method of subsection (a) we replace these two elements by new triangles, ABE and AEC to complete the solution shown (together with its projected "shadow") in Figure 7c. Although this sequence turned out to be relatively painless it was necessary to determine which element contained the projection D' of D in order to verify that D should not be reinstated. Although the element BFE was in this case not topologically very distant

from D in the original connectivity (Figure 7a) there will be cases in general where an extensive search may be needed to locate the element in which the projection D' lies.

Next, consider the alternative strategy we might adopt, eliminating vertex E first instead of D. Figure 7d shows the result of replacing the four triangles that met at E in the original configuration by the triangulation BCD, BFC needed to fill the quadrilateral gap BFCD. At this stage the structure shown in Figure 7d contains a "pocket" consisting of the pair BDC, BCD, which will mutually cancel, casting vertex D adrift. When we compare E with its image E' in ABC we find in this case that

$$R(E) < R(E')$$

and hence that E should be reinstated, replacing ABC by the new triangles, ABE, AEC and BCE. The result, illustrated in Figure 7e still contains an edge (BC) with negative dihedral which, on correction by the standard method leads at last to the final configuration illustrated in Figure (7c).

The same approach may be adopted to resolve more complicated snagged configurations than the one illustrated in Figure 7a but it is clearly not a procedure that we would wish to apply more often than we have to. In resolving snags in three dimensional solutions the reinstatement of a vertex is not significantly more complicated — it involves replacing one tetrahedron by four others. However, the task of covering the polyhedral-bounded gap left by the removal of a vertex involves rather greater complexity and almost certainly several iterative adjustments before the covering with non-inverted elements and positive-dihedral interfaces can be found. Probably the simplest gap-filling procedure is to choose one vertex from the gap-boundary and create the covering cluster of tetrahedra by linking this selected vertex to each of the triangular facets of the rest of the gap-boundary that do no already contain this vertex. Adjustments towards proper element orientation and positive dihedral-angles are then made primarily by the standard iteration and, only if necessary, by further snag-resolving iterations.

4. Initial conditions

While we have discussed how to advance a solution by small changes, the question remains: how do we obtain the initial configuration for a given problem? One solution is to construct initial conditions for the problem of interest by means of a sequence of configurations corresponding to a sequence of progressively altered problems in a chain that commences with a highly regular state whose solution is known and terminates with the initial state of the actual problem of interest. If the "default" problem at the beginning of this sequence has a solution of the generic (stable graph) type, as defined in section 2, then the same code used to adjust the connectivity in time may also be employed to adjust the connectivity in the gradual transition between default configuration and intended initial state. In the majority of the problems we might wish to investigate using the semi-geostrophic finite element model the exact positions in $\underset{\sim}{X}$-space of the individual elements at the initial time are not important — rather it is the general distribution of mass at a scale larger than the typical element-separation that is important. Thus, there is no reason in such cases why the pattern of elements should not be chosen to conform in a smoothly distorted way to the regular deployment of an ideal problem whose solution possesses a stable graph.

The simplest example of an ideal pattern for two-dimensional problems in a rectangular domain is the honeycomb arrangement of hexagonal fluid elements as illustrated for a very small problem in Figure 8a together with the corresponding triangular-grid deployment of the data in dual space, shown in Figure 8b. The corner elements are quarter-sized, the remaining elements on the boundary of the rectangular support-region of $\underset{\sim}{x}$-space are half-sized and all interior elements are complete hexagons in $\underset{\sim}{x}$-space. The solution $P(\underset{\sim}{x})$ then conforms at the vertices (except domain corners) to the quadratic form,

$$P(\underset{\sim}{x}) = \frac{1}{2}(ax^2+by^2) \ , \qquad (4.1)$$

where a and b depend on the X and Y spacing of the data elements and their mass.

The data for the initial state of interest may be a spatially distorted and non-uniformly weighted version of essentially the same elements. For example, Figure 8c shows one possible redeployment of these elements with the linking line segments now put in merely to show the original connectivity — the new connections for this problem presumably being quite different and irregular.

If X^S_α, Y^S_α, M^S_α are the coordinates and mass of element α of the default configuration and X^I_α Y^I_α M^I_α are the corresponding quantities for the intended initial configuration, then the simplest construction of the intermediate stages of the evolution from 'S' variables to 'I' variables is the linearly interpolated sequence of steps:

$$X^k_\alpha = (1 - \frac{k}{K})\, X^S_\alpha + \frac{k}{K}\, X^I_\alpha \quad , \quad k = 0, 1 \ldots K \qquad (4.2)$$

and similarly for Y^k_α and M^k_α. Other prescriptions for the evolution might be more appropriate when the distortion required is fairly severe.

In the case of three-dimensional solutions the simplest pattern of data that leads naturally to a configuration of elements of the generic form (i.e., fluid or boundary elements meeting four at a time at each vertex) is that comprising a "body centred cubic" structure (e.g., see Ziman, 1972). In this structure the data points in dual space can be thought of as each belonging to either the "even" rectangular lattice of points,

$$\underline{X} = (i\Delta X \, , \, j\Delta Y \, , \, k\Delta Z), \qquad (4.3)$$

where i, j and k are all even integers, or else to the "odd" lattice of points where i, j and k are all odd integers. Each fluid element then touches eight nearest neighbours that belong to the alternate lattice, and six more distant neighbours belonging to the same lattice. Thus the

elements are 14-sided polyhedra (truncated octahedra) except for those at the boundary which are reduced to half, a quarter or even an eighth of their full size according to whether they occupy one, two or three boundary-planes of the domain.

Figure 9 sketches the arrangement of the bottom half-layer (part of the even-lattice) together with a view of the odd-lattice element closest to the corner as it would appear raised above the layer below. The connectivity of the default problem vertices in physical space, while much more complicated than in two dimensions, does nevertheless possess a very regular pattern that allows automatic construction of a table such as that depicted in Figure 1a. Thus, three-dimensional initial state configurations can be evolved in gradual stages from a known stable default configuration as in the two-dimensional cases.

Although the examples illustrated are of rectangular or cuboidal domains, it is equally possible to use the evolutionary technique to initialize solutions requiring single or double periodicity since the default patterns are already in a form compatible with assumptions of periodicity.

5. Concluding remarks

This note has set out to define a broad strategy by which the finite element (geometric) models of semi-geostrophic theory can be efficiently implemented to solve time dependent problems in bounded or periodic domains. The necessity in such a scheme to refer frequently to all the extries of large tables of vertex and element statistics and to alter these tables piece by piece requires a data-structure with considerable versatility. Look-up efficiency implies the need for a tree-like hierarchy associated with the tables in order bring access time within order Log N, while the need to update frequently the table entries requires that the hierarchical structure is of an adaptable form.

In view of the significant investment in skilled programming that would be required to develop the necessary software to implement an efficient version of the panel-beater algorithm it is pertinent to consider carefully the kind of problems that could be tackled within the geometric model framework that could not easily be dealt with by more conventional numerical methods. The distinctive advantage of the geometric method is, of course, its complete immunity to adverse effects of solution discontinuities which cause considerable problems in finite difference or spectral methods. It is therefore appropriate that the problems to which we might devote our attention are those in which the presence of a discontinuity, in the form of a definite front or inversion, is an essential feature. As noted in the introduction, the development of a mature baroclinic wave through the latter stages of its life-cycle falls into this category and is a case in which the geometric model, given sufficient resolution, could supply important insights. Another example would be the examination of the natural baroclinic modes that can propagate on an existing frontal interface of realistic structure. Convective solutions, even of the penetrative kind studied by Shutts and Cullen (1987), could be modelled by the panel-beater algorithm provided measures are taken to slow the movement of convecting elements from discontinuous jumps to more manageable continuously ascending migrations (e.g. by lagging the effects of latent heat release.)

The panel-beater algorithm may enjoy advantages over the more "brute-force" methods of solving the geometric model problem when it comes to generalising semi-geostrophic theory to accommodate spherically curved domains or variable Coriolis parameter where the hypersurfaces of the fluid elements can only locally be represented as Euclidean hyperplanes. Because the essential operations of the panel-beater algorithm are local, the restrictions of more general forms of the semi-geostrophic theory, such as Salmons' (1983, 1985) extensions are unlikely to prove a serious hindrance.

In a domain of variable Coriolis parameter the construction of solutions using an envelope of hyperplanes must be abandoned. It is argued in Purser (1988) that the natural and simplest generalisation of the

geometric model in such a case, and one that turns out to be consistent with Salmon's approach, is to consider the solution instead as the envelope of a collection of intersecting concave paraboloidal functions instead, with the peak-curvature of each paraboloid determined by the local Coriolis parameter. Thus, although the edges and intersections of cells of the geometric solution are now curved, in the immediate vicinity of a topological transition the arrangement still may be given an Euclidean dual-space interpretation with relative orientations obtained from the tangents of the edges involved near the locus of the transition in physical space.

References

Chynoweth, S. 1987 'The semi-geostrophic equations and the Legendre transformation'. Ph.D. Thesis, Dept. of Mathematics, Univ. of Reading, England.

Cullen, M. J. P. 1983 'Solutions to a model of a front forced by deformation'. Quart. J. Roy. Meteor. Soc., 109, 565-573.

Cullen, M. J. P. and Purser, R. J. 1984 'An extended Langrangian theory of semi-geostrophic frontogenesis. J. Atmos. Sci., 41, 1477-1497.

Cullen, M. J. P., Chynoweth, S. and Purser, R. J. 1987a On semi-geostrophic flow over synoptic-scale topography. Quant. J. Roy. Meteor. Soc., 113, 163-180.

Cullen, M. J. P., Norbury, J., Purser, R.J. and Shutts, G. T. 1987b 'Modelling the quasi-equilibrium dynamics of the atmosphere'. Quart. J. Roy. Meteor. Soc., 113, 735-758.

Gilmore, R. 1981 'Catastrophy theory for scientists and engineers'. Wiley & Sons, 666 pp.

Hoskins, B. J. 1975 'The geostrophic momentum approximation and the semi-geostrophic equations'. J. Atmos. Sci., 32, 233-242.

Hoskins, B. J. and Bretherton, F. P. 1972 Atmospheric frontogenesis models: Mathematical formulation and solutions'. J. Atmos. Sci., 29, 11-37.

Purser, R. J. and Cullen, M. J. P. 1987 'A duality principle in semi-geostrophic theory'. J. Atmos. Sci., 44, 3449-3468.

Purser, R. J. 1988 'Variational aspects of semi-geostrophic theory'. UK Met. Office Met O 11 Scientific Note No. 5.

Rockafellar, R. T. 1970 'Convex analysis'. Princeton University Press, 451 pp.

Salmon, R. 1983 'Practical use of Hamilton's principle'. J. Fluid Mechs., 132, 431-444.

Salmon, R. 1985 New equations for nearly geostrophic flow', J. Fluid Mechs., 153, 461-477.

Shutts, G. J. and Cullen, M. J. P. 1987 'Parcel stability and its relation to semi-geostrophic theory'. J. Atmos. Sci., 49, 1318-1330.

Shutts, G. J., Booth, M. and Norbury, J. 1988 'A geometric model of balanced, axisymmetric flows with embedded penetrative convection'. J. Atmos. Sci., 45 (to appear).

Ziman, J. M. 1972 'Principle of the theory of solids' (second edition). Cambridge University Press 435 pp.

Figure Captions

Figure 1: (a) A possible format for the tabulation of vertex data corresponding to the portion of the configuration illustrated in (b). The Chiral signature in (a) informs whether the adjacent cells taken in the order given form a right-handed (+) or left-handed (-) set.

Figure 2: (a) Schematic view of a polyhedral solution and (b) its dual.

Figure 3: Cross-section through a polyhedral solution in the x-P plane illustrating the possibility of "closing" the polyhedral surface with a "lid".

Figure 4: Sequence illustrating a first order transition for a two-dimensional solution, (a) before; (c) at and (e) after the transition in physical space correspond to (b), (d) and (f) respectively in dual space.

Figure 5: Transition associated with the shrinking and extinction of a triangular element. (a) Initial state; (b) solution dual to (a); (c) effect of panel-beater replacement of edge AB with edge CD; (d) final solution in physical space following pair-annihilation at +ACD, -ACD; (e) corresponding dual space result.

Figure 6: Schematic representation of typical transition in a three-dimensional solution as viewed in X, Y, Z-space. (a) the polyhedron formed from either (b) two simplices meting at interface ABC or (c) three simplices meeting at edge DE. A typical panel-beater iteration changes the composition (b) to (c) or vice-versa.

Figure 7: (a) Portion of a two-dimensional snagged solution and its shadow projected on the XY-plane. Note that all edges possess positive dihedral angle when the orientation of surface facets is taken into account. (b) Result of removing vertex D from the solution. The new element BCE, although inverted, is removed (c) by the regular panel-beater process acting to replace negative dihedral-angle at edge BC by positive dihedral edge AE. (d) An alternative snag-removing strategy begins with the elimination (temporarily) of vertex E leaving "pocket" BCD; (e) reinstating E and applying pair annihilation to remove the pocket still leaves negatively angled edge BC which is replaced by new edge EF in a regular panel-beater iteration to give the result shown in (c).

Figure 8: (a) Pattern in $\underset{\sim}{x}$-space for a standard "default" problem in two dimensions. (b) Corresponding pattern in $\underset{\sim}{X}$-space. (c) A smoothly deformed version of the pattern (b) that might constitute initial data for a problem of interest. Dashed lines show the original connectivity.

Figure 9: Schematic view of lowest half-layer of $\underset{\sim}{x}$-space elements of default problem in three-dimensions, together with view of a complete element lifted from its positon near the domain corner. Complete elements are 14-sided truncated octahedra which pack to give a body-centred-cubic arrangement consistent with the requirement for a stable graph.

VERTEX	ADJACENT CELLS			NEIGHBOURING VERTICES			CHIRAL SIGN.
1	101	103	104	5	2	10	+
2	101	102	104	3	1	6	−
3	102	104	105	4	7	2	+
4	104	105	106	8	5	3	−
5	103	104	106	4	9	1	−

FIGURE 1a

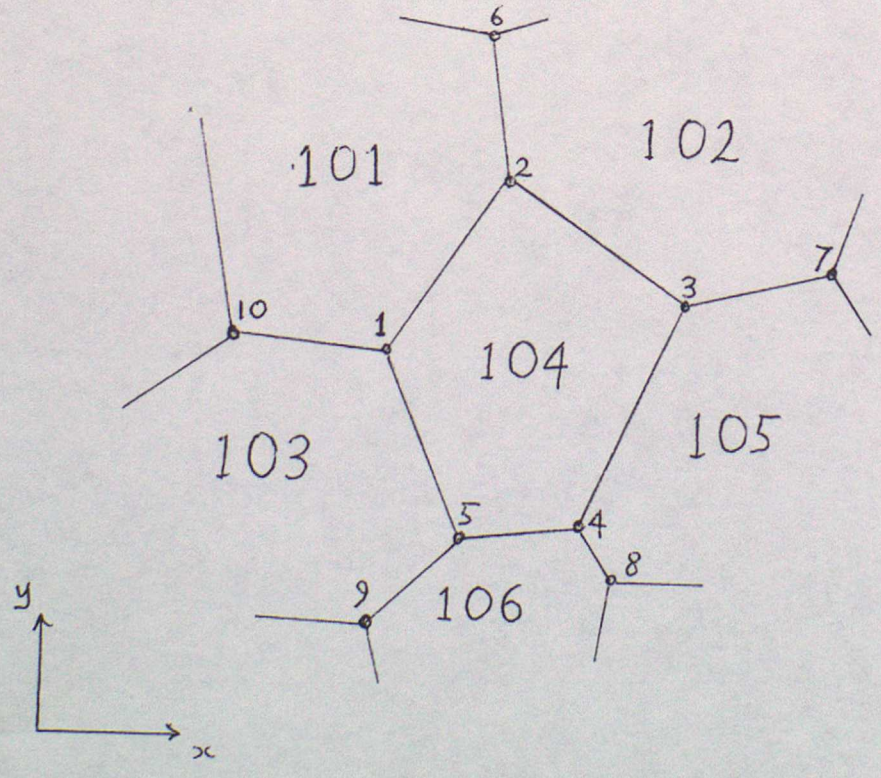
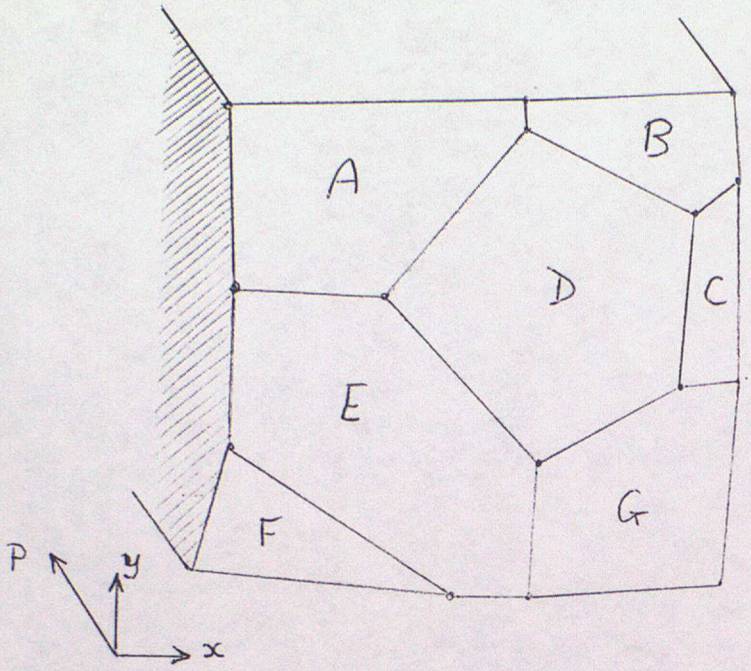
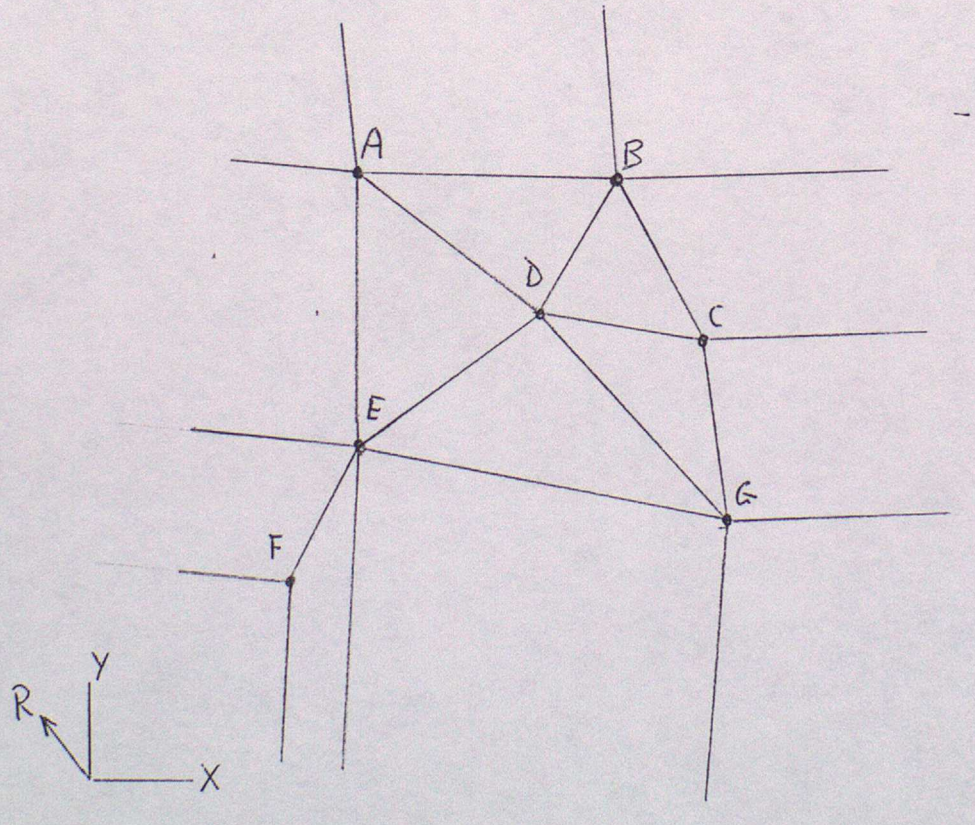


FIGURE 1b



(a)



(b)

FIGURE 2

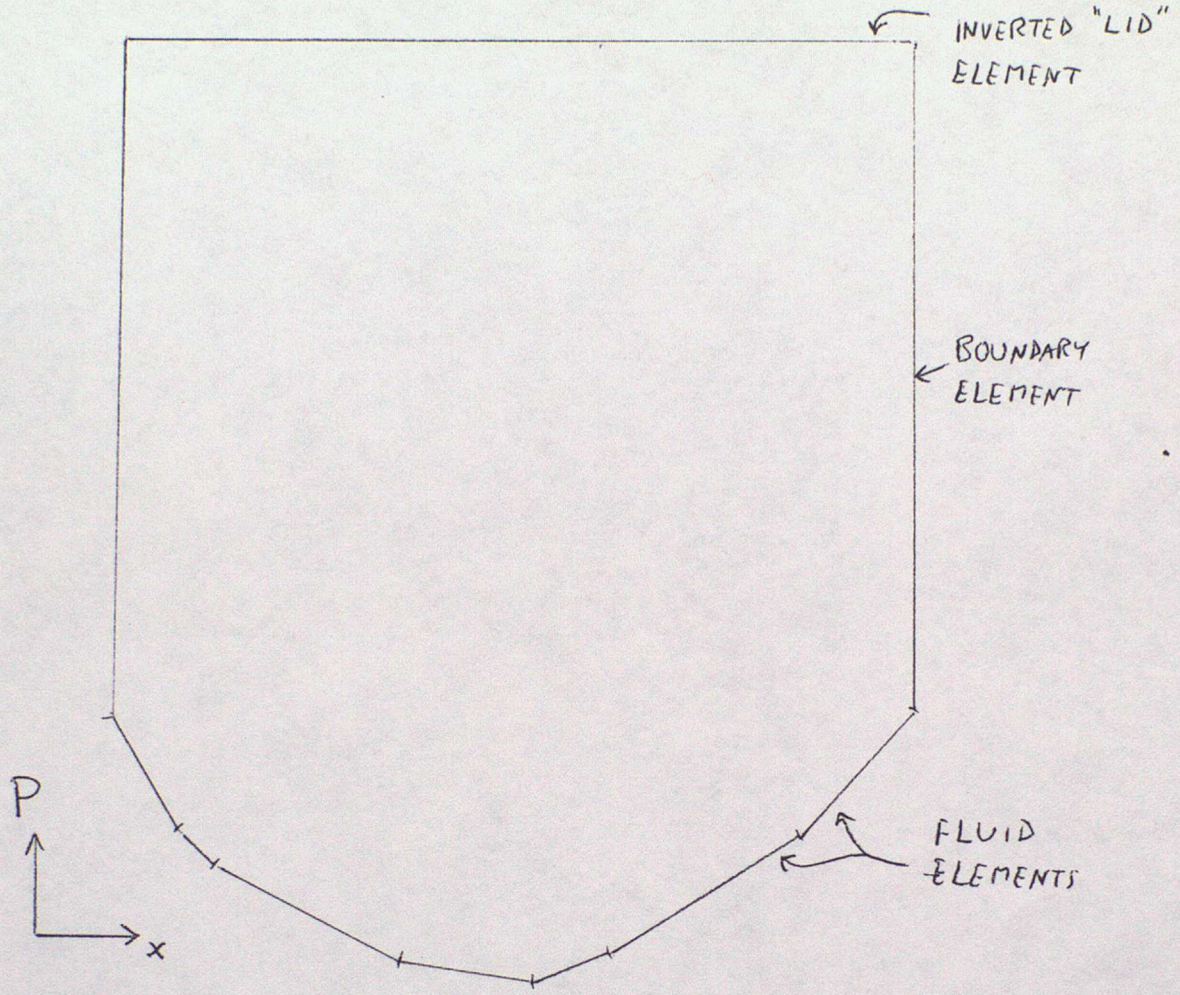
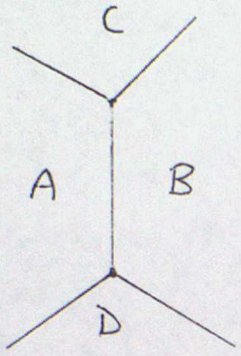
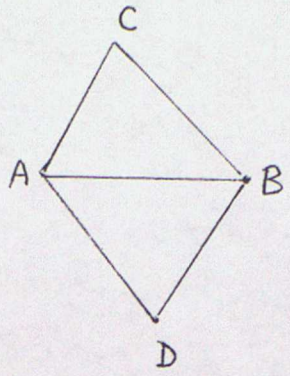


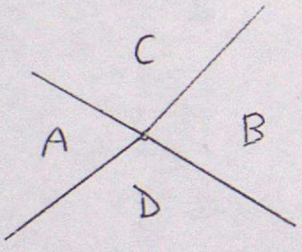
FIGURE 3



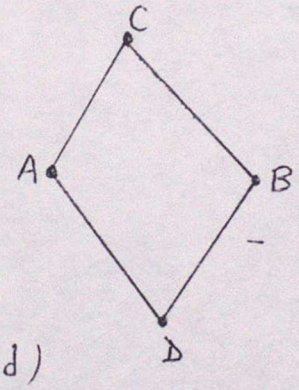
(a)



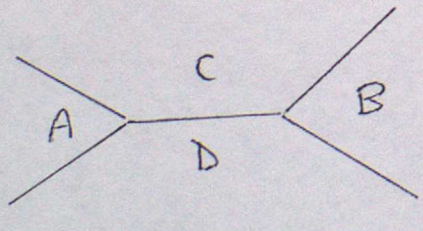
(b)



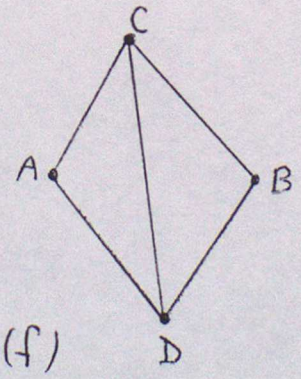
(c)



(d)



(e)



(f)

FIGURE 4

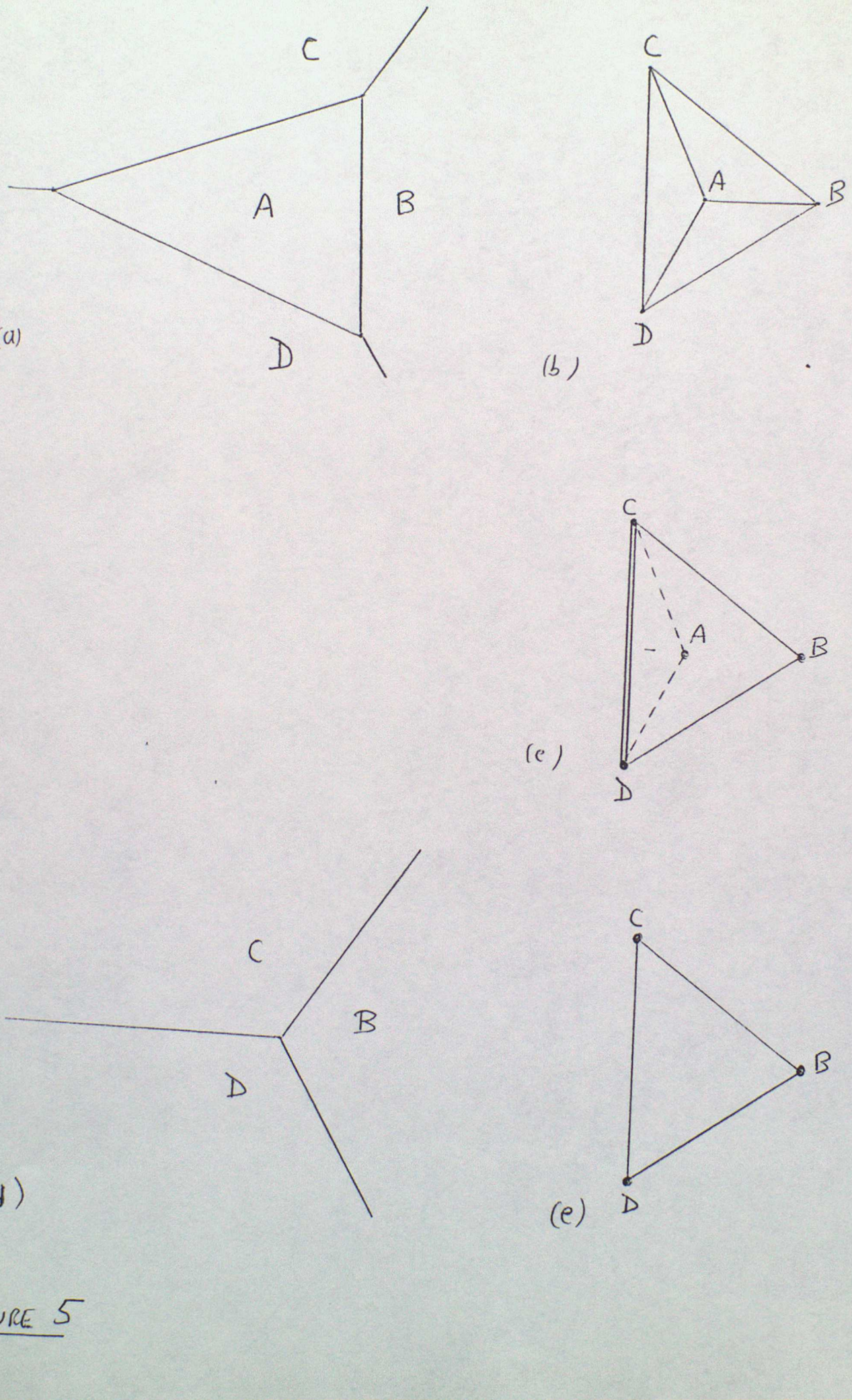
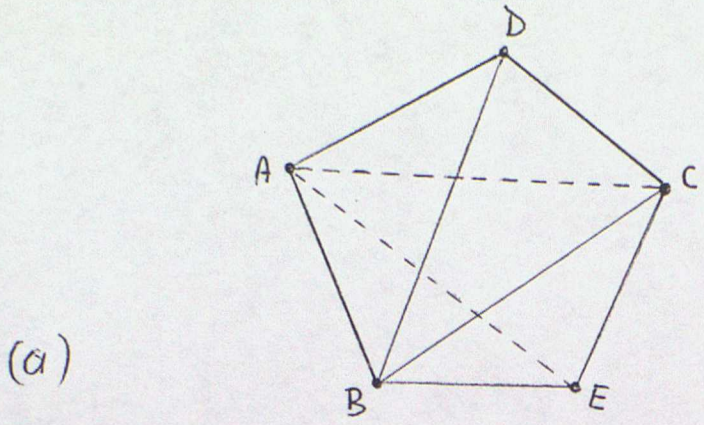
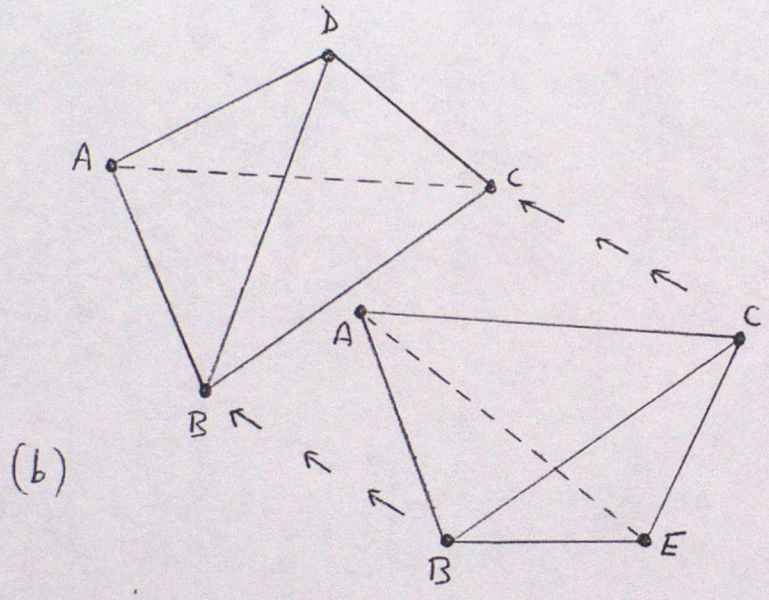


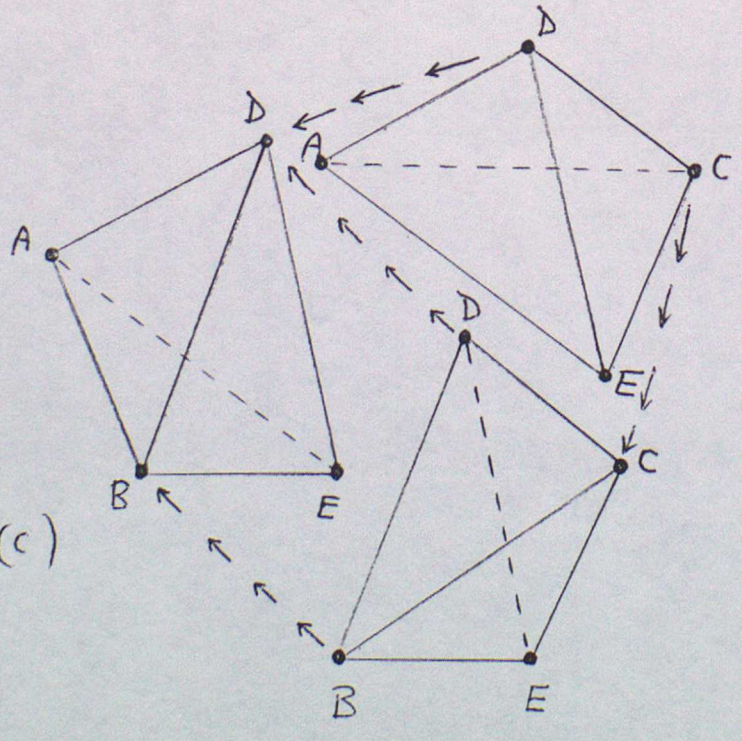
FIGURE 5



(a)



(b)



(c)

FIGURE 6

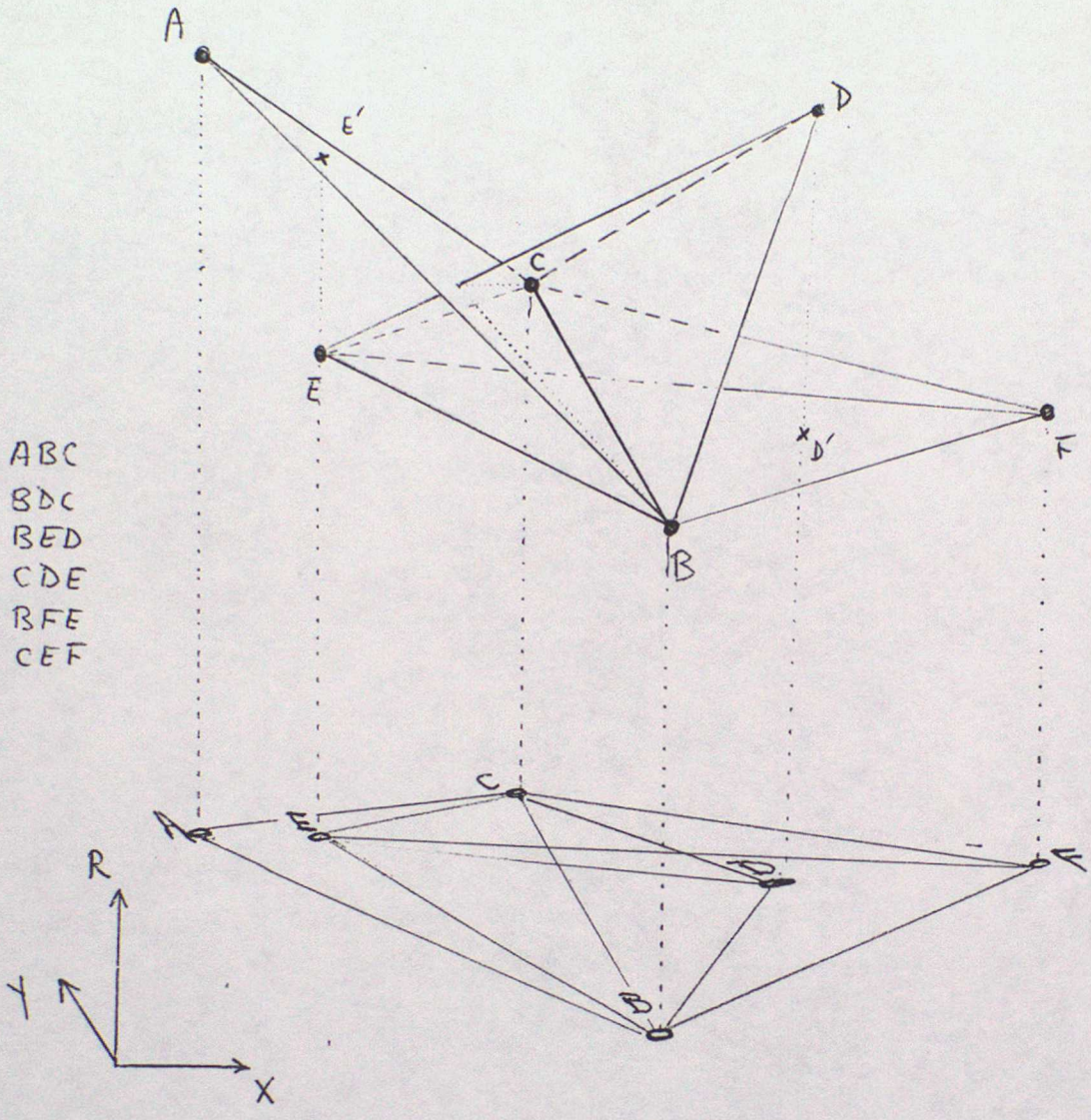
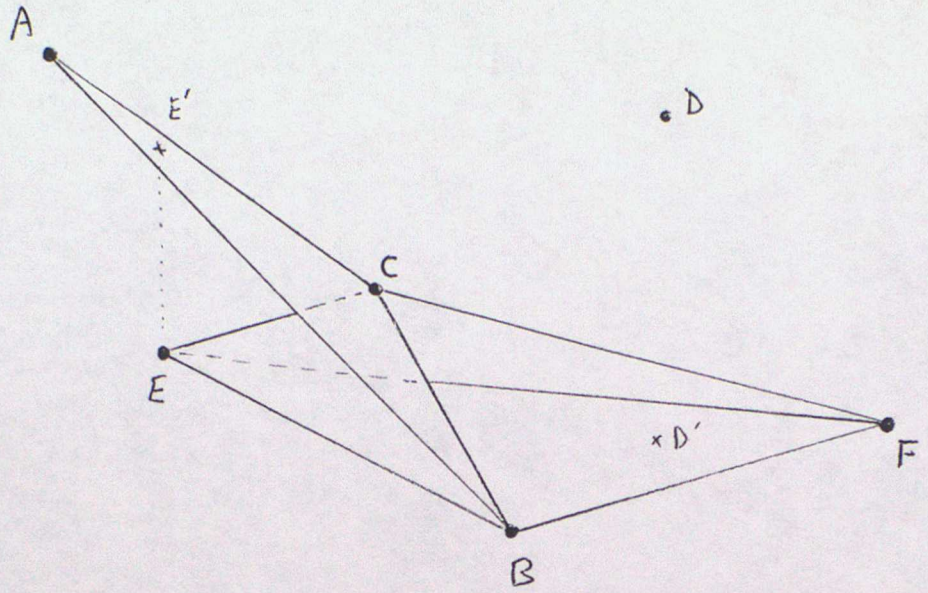
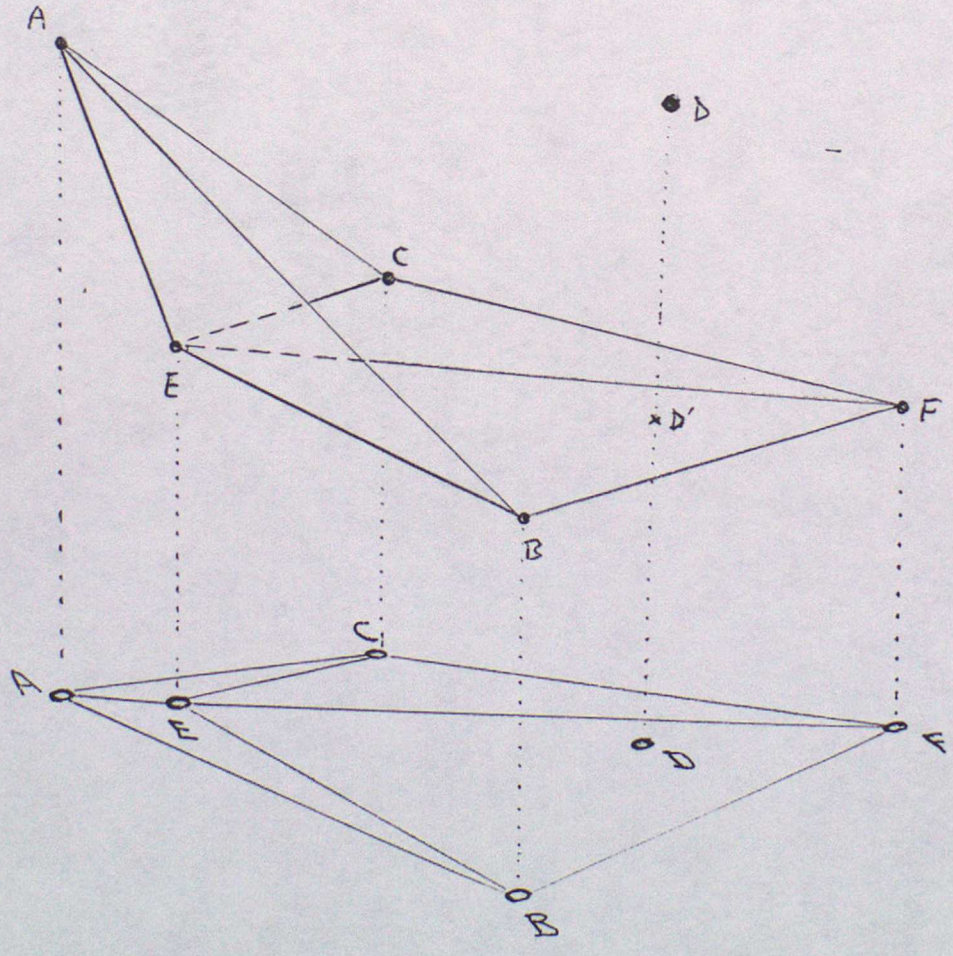


FIGURE 7a



ABC
BEC
BFE
CEF

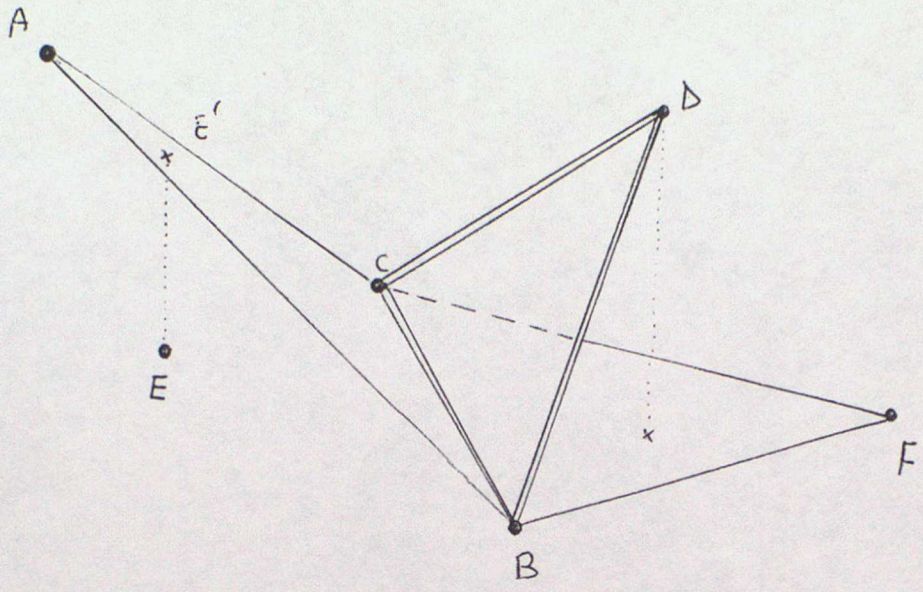
(b)



ABE
AEC
BFE
CEF

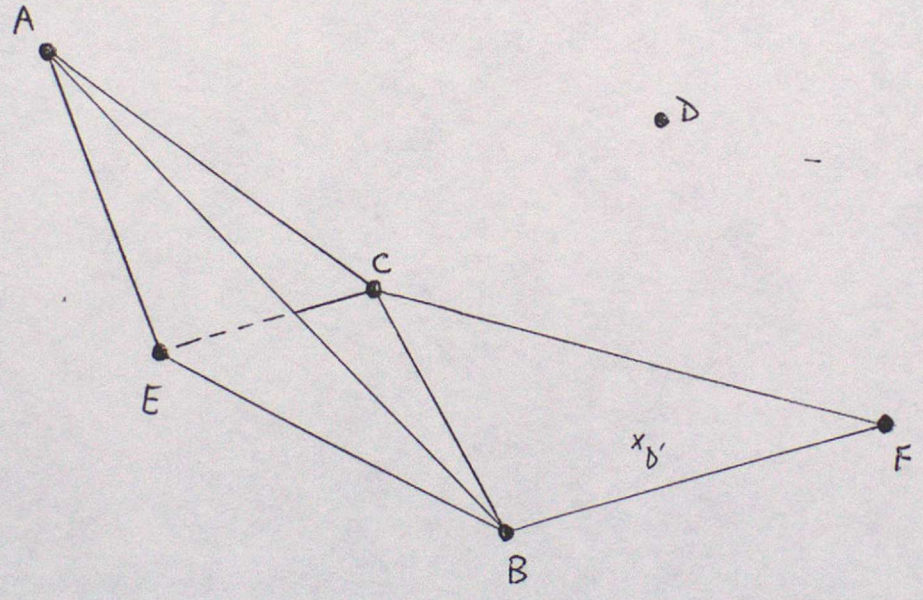
(c)

FIGURE 7



ABC
BDC
BCD
BFC

(d)



ABE
AEC
BCE
BFC

(e)

FIGURE 7

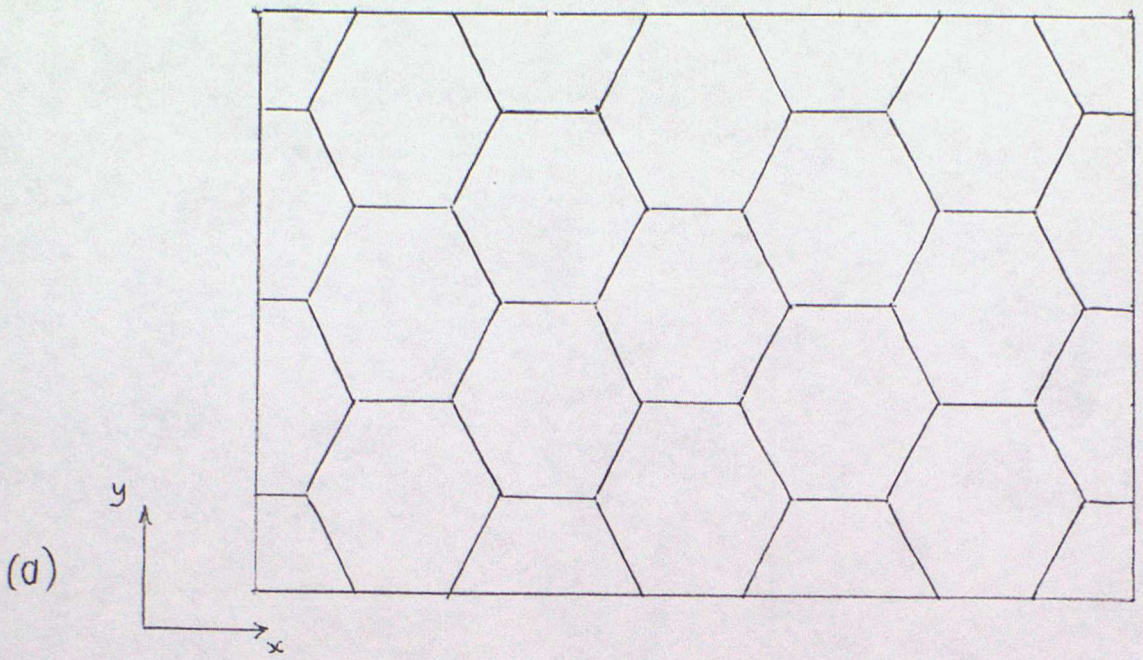


FIGURE 8

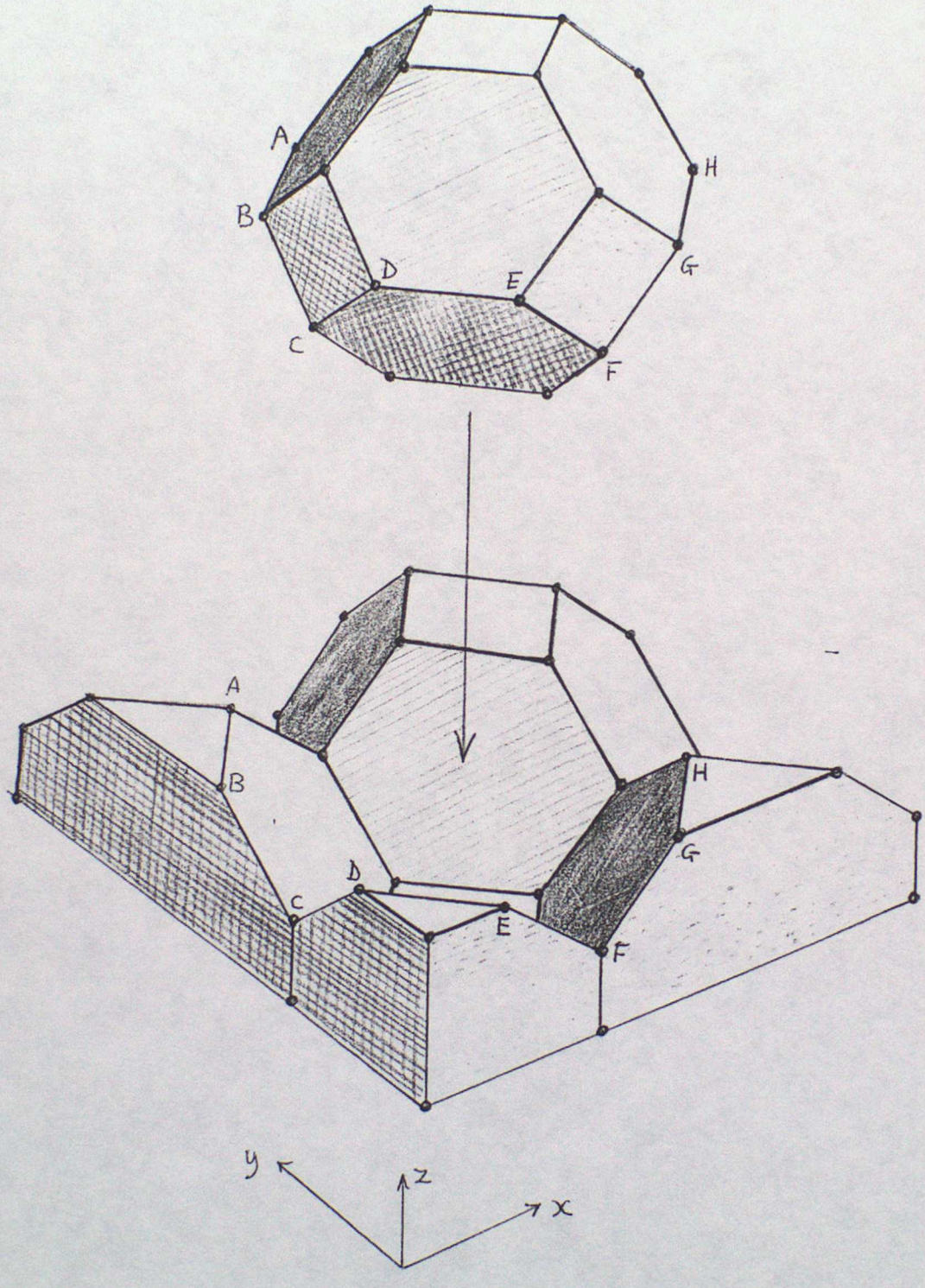


FIGURE 9