Chapter V

Informational Efficiency of Mechanisms

An important motivation for developing the (r,d)-network model of computing is to use it to analyze the computational tasks carried out by economic mechanisms. In particular we wish to study the tradeoffs, if any, between the communications requirements and the computational requirements of achieving a given economic performance by a decentralized mechanism. There are, of course many different mechanisms and for each many computational tasks that could be studied. A particular case is that of a static decentralized mechanism that realizes the Walrasian performance function. In Chapter VII we apply the (r,d)-network model to analyze an example of this kind, a two person two good exchange economy. In this chapter we provide background for the analysis carried out in Chapter VII, and define some concepts needed in order to make the model applicable to that and similar examples.

The general set-up studied is as follows. There are n agents, 1,...,n. Each agent has environmental characteristics denoted e^i; the set of possible environments for agent i is E^i. The joint environment $e = (e^1,...,e^n)$, is by assumption an element of

$E = E^1 x \ldots x E^n$. It is also assumed that agent i initially knows his characteristic e^i, and that is all he/she knows directly about the joint environment e.

Let A denote the space of joint actions or outcomes. In the case of an exchange environment these are trades or allocations. There is a function F:E --> A which expresses the goals of economic activity. In our example F(e) is the (unique) Walrasian trade when the environment is $e \in E$.

We consider mechanisms

$$\pi = (\mu, M, h)$$

where

$$\mu : E \dashrightarrow M$$

is a privacy preserving correspondence, called the <u>message correspondence</u>, M is the <u>message space</u> of the mechanism, and

$$h : M \dashrightarrow A$$

is a function with the property that h is constant on the sets μ(e) for all e in E. The function h is the <u>outcome function</u> of the mechanism. The mechanism π <u>realizes</u> F on E if for all $e \in E$

$$h(\mu(e)) = F(e).$$

The message correspondence μ is <u>privacy preserving</u> if for each i=1,...,n, there exist correspondences

$$\mu^i : E^i \dashrightarrow M$$

such that

$$\mu(e) = \cap_i \mu^i(e^i).$$

The requirement that μ preserve privacy is that the message of an agent can depend only on that agent's environmental component and on the messages received from other agents.

Such a mechanism π can be given directly, or can be regarded as the equilibrium form of a dynamic message exchange process in which the agents exchange messages taken from the space

$$M = M^1 x \ldots x M^n$$

according to prescribed rules

$$f^i : M \times E^i \rightarrow M^i,$$

where,

$$f^i(m(t), e^i) = m^i(t+1),$$

for $i=1,\ldots,n$ and $t=1,2,\ldots,$. The initial message $m(0)$ is given.

(Here privacy preserving is a property of the functions f^i.)

The stationary messages defined by this system of difference equations are given by

$$0 = g^i(m, e^i) = f^i(m, e^i) - m^i,$$

for all $i=1,\ldots,n$.

We define

$$\mu^i(e^i) = \{m \in M \mid g^i(m, e^i) = 0\}.$$

We shall focus attention on mechanisms in equilibrium form. Even abstracting from the dynamics of message exchange several different computational tasks can be distinguished. One interpretation of decentralized mechanisms in equilibrium form is the verification scenario. In this scenario, a candidate equilibrium message $m \in M$ is 'posted', and seen by each agent. Each agent i separately checks the message to see whether it satisfies his equilibrium condition. If it does, agent i says "Yes", if not, he says "No". If all agents say "Yes" to a given message, then it is verified to be an equilibrium message. That is, there are individual verifier functions, V^i, for $i=1,\ldots,n$

$$V^i(m,e^i) = \begin{cases} 1 & \text{if } g^i(m,e^i)=0 \\ 0 & \text{otherwise,} \end{cases}$$

and a verification function

$$V : \{0, 1\}^n \to [0,1]$$

given by

$$V(x)=(1/n) \Sigma x^i,$$

where

$x^i = V^i(m,e^i)$, for $i = 1, \ldots, n$, and $x=(x^1, \ldots, x^n)$.

The computational tasks involved in this are:

(i) to determine whether $g^i(m,e^i)=0$, for each i, given m,

(ii) to evaluate V, and

(iii) to evaluate h.

Presumably the V^i's are computed by the individual agents, and the function V by some institution, perhaps personified by an additional agent. In this scenario the origin of the 'posted' message is not considered, nor are the verifying messages, (the values of V^i) counted in the message space.

Another interpretation is that each agent i transmits the subset $\mu^i(e^i)$ to a central institution that finds the equilibrium, e.g., clears the market. Finding equilibrium is most naturally addressed in a dynamic setting, but since much of the research on message space size has been done in the context of equilibrium mechanisms, and since it is our objective to illustrate the application of the (r,d)-network model to mechanisms, it is not unnatural to begin by studying tradeoffs between communication and computational complexity in that setting. Thus, we adopt the second interpretation of the equilibrium model, one in which the equilibrium is computed from the individual message correspondences. This may be thought of as an iterative dynamic process that finds the equilibrium in one step.

In this interpretation, the computational task is to compute the set $\mu(e)$ from the sets E^i, and to evaluate the outcome function h. If we are to model this computation by (r,d)-networks, we must confront the fact that inputs to such a network must be d-dimensional Euclidean vectors. In Chapter IV we have given the definition of a network that computes an encoded version of a function. The computational task is then to compute an encoded version of the set $\mu(e)$ and to compute an encoded version of the function h.

Assumption 5.1.

The set M is a manifold of dimension p, and the sets E^i are manifolds of dimension q^i, so that E is a manifold of dimension $q = \Sigma q^i$.

The computation of the equilibrium message correspondence and of the outcome function are related. By changing coordinates in the message space it is possible to shift the burden of computation between them. We make the following simplifying assumption on the mechanisms considered, in effect combining these two tasks.

Assumption 5.2

(i) The message correspondence m is privacy preserving and single valued.

(ii) There is a p_1 dimensional submanifold M_1 of M, such that h is a projection onto M_1.

We restrict attention to mechanisms satisfying Assumptions 5.1 and 5.2. Given a goal or performance standard F:E-->A, we may consider the class of mechanisms that realize F. For each such mechanism there are two indicators or measures of informational requirements, namely, the dimension, m, of the message space M of the mechanism, and the time, t, required to compute the equilibrium message $\mu(e)$ in M. By Assumption 5.2 the time to compute the outcome function is already incorporated in the computation of $\mu(e)$. Thus, each mechanism realizing F and satisfying Assumptions 5.1 and 5.2 has associated to it a point, (m,t), (with integer coordinates) in R^2. We may refer to the set of points so defined as the _informational image_ of the set of mechanisms realizing F and satisfying Assumptions 5.1 and 5.2. The _efficient frontier_ of this informational image describes the available tradeoffs between communication and computation in the realization of F. In Chapters VII and VIII we apply the (r,d)-network model, with r=2 and d=1, and with the modules required to be analytic, to find the efficient frontier of the class of mechanisms that realize the Walrasian performance standard on the class of two person two good exchange environments presented there.

Chapter VI

Essential Revelation Mechanisms, Differentiably Separable Functions and the Theorems of Leontief and Abelson

In this chapter we discuss the relation between a generalization, due to Abelson [1], of a result of Leontief [15] and a type of mechanism called an adequate revelation mechanism. Suppose that a network computes an encoded version of a function G, where the encoding of the range of G is given by functions $\{h_j; 1 \le j \le t\}$. Suppose that S(i;j) is an LE-i-separator set for the j^{th} output vertex of the network where the j^{th} output vertex is associated to the function h_j. The concept of LE-i-separator set was introduced in Chapter IV. When the spaces $(X_i/h_j \cdot G)$ are Hausdorff, around each point s in S(i;j) there is a neighborhood U_s such that the restriction of q_i to U_s is a homeomorphism from U_s to a subspace $V(U_s)$ of $(X_i/h_j \cdot G)$. If the spaces $(X_i/h_j \cdot G)$ are manifolds, then this gives an upper bound on the dimension of separator sets. In the first section of this chapter we give conditions on a real valued function F that guarantee that if the quotient space (X_i/F) is Hausdorff, then (X_i/F) has the structure of a topological manifold.

The conditions are rank conditions on a submatrix of the Hessian of F. These rank conditions are used by Leontief [15] to study production functions and by Abelson [1] to study the minimum communication requirements of a distributed computation. In the second section we discuss the concept of adequate revelation mechanism and its relation to the (X_i/F). When the spaces (X_i/F) are manifolds then, under suitable global conditions, it is possible to characterize the space $(X_1/F) \times \ldots \times (X_n/F)$ as a "smallest" message space for a mechanism whose message space is a product of individual messages spaces, one space for each agent. More precisely, for each mechanism whose message space is a product of individual message spaces $M_1 \times \ldots \times M_n$ with a message correspondence

$$\mu^1 \times \ldots \times \mu^n : X_1 \times \ldots \times X_n \longrightarrow M_1 \times \ldots \times M_n,$$

there is a function $g_1 \times \ldots \times g_n$, such that $g_i \cdot \mu^i = q_i$, where q_i is the quotient map from X_i to (X_i/F). Loosely speaking, the quotient map q_i squeezes out as many variables as possible.

Section I.

The Theorems of Leontief and of Abelson

In this section we introduce the notation and the results needed to explain the relation between a more general form of Leontief's theorem and adequate revelation mechanisms. A relation between these two concepts involves the concept of differentiable separability. Differentiable separability also plays an important role in Chapter X, where we analyze the relation between the Dimension Based Lower Bound on the time required to compute an encoded version of a function F and a bound on the time required for finite networks to compute approximations to the function F.

Suppose that $F(x_1,...,x_N)$ is a function of N variables. If $\alpha=(\alpha(1),...,\alpha(N))$ is a sequence of nonnegative integers, denote by $|\alpha|$ the sum $\alpha(1)+...+\alpha(N)$. If F has continuous partial derivatives to order $d\geq\alpha$, then denote by

$$D(x_1^{\alpha(1)}....x_N^{\alpha(N)};F)$$

the derivative

$$\partial^{|\alpha|}F/\partial x_1^{\alpha(1)}...\partial x_N^{\alpha(N)}.$$

Suppose that $E^1,...,E^n$, are Euclidean spaces of dimensions $d(1),...,d(n)$, respectively. We suppose that the space E^i, $1\leq i\leq n$ has coordinates $x_i=\{x_{i\,1},...,x_{i\,d(i)}\}$. Assume that $(p_1,...,p_n)$ is a point of $E^1x...xE^n$, and assume that U_i is an open neighborhood of the point p_i for $1\leq i\leq n$. Suppose that F is a real valued C^2-function defined on $U_1x...xU_n$. We introduce two matrices in (I) and (II), below.

(I): The matrix

$$BH(F:x_{i\ 1},\ldots,x_{i\ d(i)};x_{1\ 1},\ldots,x_{i-1\ d(i-1)},$$
$$,x_{i+1\ 1},\ldots,x_{n\ d(n)})=$$
$$BH(F:x_i;x_{<-i>})$$

is a matrix that has rows indexed by

$$x_{i\ 1},\ldots,x_{i\ d(i)}$$

and columns indexed by

$$F,x_{1\ 1},\ldots,x_{i-1\ d(i-1)},x_{i+1\ 1},\ldots,x_{n\ d(n)}.$$

The entry in the $x_{(i\ u)}{}^{th}$ row and in the F column is $D(x_{i\ u};F)=\partial F/\partial x_{i\ u}$. The entry in row $x_{i\ u}$ and in column $x_{j\ w}$ is

$$D(x_{i\ u}\ x_{j\ w};F)=\partial^2 F/\partial x_{i\ u}\ \partial x_{j\ w}.$$

The matrix $BH(F:x_i;x_{<-i>})$ is a type of bordered Hessian because it consists of a matrix of second derivatives bordered by collection of columns of first derivatives.

(II):

The matrix

$$H(F:x_i;x_{<-i>})$$

is the submatrix of $BH(F:x_i;x_{<-i>})$ that consists of the columns indexed by $x_{u\ v}$, $u\in\{1,\ldots,i-1,i+1,\ldots,n\}$ and $1\le v\le d(u)$. In other words, we derive H from BH by eliminating the column indexed by the function F.

In case that the number of Euclidean spaces is two, so that $F:E^1 \times E^2 \longrightarrow R$, we use a slightly less cumbersome notation. Suppose that E^1 has coordinates $\{x_1,\ldots,x_p\}$ and E^2 has coordinates $\{y_1,\ldots,y_q\}$, then we use as row indices for $BH(F:x_1,\ldots,x_p;y_1,\ldots,y_q)$ the variables $x_1,\ldots,x_p$ and as column indices F, $y_1,\ldots,y_q$. The $(x_i,F)^{th}$ entry in $BH(F:x_1,\ldots,x_p;y_1,\ldots,y_q)$ is

$$\partial F/\partial x_i = D(x_i;F)$$

and the $(x_i,y_j)^{th}$ entry is

$$D(x_i\ y_j;F)=\partial^2/\partial x_i\ \partial y_j.$$

The matrices $BH(F:x_i;x_{<-i>})$ and $H(F:x_i;x_{<-i>})$ are matrices of functions in the coordinates $x_1,\ldots,x_n$ of $E^1x\ldots xE^n$. The conditions we place on the matrices BH and H require that some, but not all, of the variables are to be evaluated at a point. When that partial evaluation takes place we indicate this by adding an asterisk to the H or BH. Specifically,

(III): The matrix

$$BH^*(F:x_i;x_{<-i>})[\ x_i,p_{<-i>}\ ]$$

is the matrix that results from evaluating the variables

$$x_1,\ldots x_{i-1},x_{i+1},\ldots x_n$$

of the entries of $BH(F:x_i;x_{<-i>})$ at the point $p_{<-i>}=(p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_n)$. The matrix $BH^*(F;x_i,x_{<-i>})[\ x_i,p_{<-i>}\ ]$ is a function of the variables $x_{i\ 1},\ldots,x_{i\ d(i)}$ alone. Similarly, the matrix

$$H^*(F:x_i;x_{<-i>})[\ x_i,p_{<-i>}\ ]$$

is the submatrix of $BH^*(F:x_i;x_{<-i>})[\ x_i,p_{<-i>}\ ]$ derived by deleting the column indexed by F.

If a continuous (r,1)-network can compute a function $F(\ x_1,\ldots,x_m;y_1,\ldots,y_n\ )$ in two units of time then, as we have seen in Chapter III, the function F can be written as a superposition C(A,B) where each of A and B is a function of at most r variables. Lemma 6.1 establishes a criterion to decide if F can be computed by an (r,1)-network when F is sufficiently differentiable. The criterion is given in terms of the matrix BH(F:x;y).

<u>Lemma 6.1</u>. Suppose that

(i) $F(\ y_1\ ,\ldots,y_m;x_1,\ldots,x_n\ )=$

$C(\ y_1\ ,\ldots,y_m\ ;A_1,\ldots,A_r\ )$,

where C is a function of m+r variables with continuous d^{th} derivatives $d\geq 2$,

(ii) each A_i is a function of n variables $\{x_j;1\leq j\leq n\}$ that has continuous d^{th} derivatives.

Then, BH(F:x;y) has rank less than or equal to r.

Proof. The Chain Rule shows that

$D(x_i;F)=\Sigma_k\ \ D(A_k;C)\ D(x_i;A_k)$,

and therefore

$D(x_iy_j;F)=\Sigma_k\ D(A_ky_j;C)D(x_i;A_k)$.

The matrix BH is the product of the matrix

($D(A_k y_j;C)$), which has at most r linearly independent columns, and the matrix ($D(x_i;A_k)$). Therefore, BH has rank at most r.▒

More generally the following statement is easy to prove.

Theorem 6.1. Suppose that F is a function of N= d(1)+...+d(r) real variables

$x_{1\ 1},\ldots,x_{1\ d(1)};x_{2\ 1},\ldots;x_{r\ 1},\ldots,x_{r\ d(r)},$

where $d(i)\geq 1$ for each $1\leq i\leq r$.

(1) Denote by $TBH_i(\underline{F}:x_i;x_{<-i>})$ the infinite matrix that has rows indexed by the variables

$$x_{i\ 1},\ldots,x_{i\ d(i)},$$

and columns indexed by F and the monomials

$x_{1\ 1}^{\alpha(1\ 1)}\ldots x_{j-1\ d(i-1)}^{\alpha(i-1)\ d(i-1))}x_{i+11}^{\alpha(i+11)}\ldots x_{rd(r)}^{\alpha(rd(r))}$

(that is, the exponents are $\alpha(1\ 1),\ldots,\alpha(r\ d(r))$ with $\alpha(j,k)=0,\ 1\leq k\leq d(j)$), such that $D(x_{i\ k};F)$ is in the $(x_{i\ k},F)^{th}$ position and $D(x_{i\ k}M;F)$ is the entry in the position with row index $x_{i\ k}$ and column index the monomial M,

(2) If x* is an N dimensional vector of real

numbers, denote by $TBH_i(F)*(x*)$ the matrix $TBH_i(F)$ with each entry evaluated at the vector x*.

Then a necessary condition that there are functions

$A_1(x_{1\ 1},\ldots,x_{1\ d(1)})$,

$\ldots,A_r(x_{r\ 1},\ldots,x_{r\ d(r)})$

and

$C(y_1,\ldots,y_r)$,

where

(a) each A_j is defined in a neighborhood U of x*,

(b) C is defined in a neighborhood V of $(A_1(x*),\ldots,A_r(x*))$ that contains the set $(A_1(x),\ldots,A_r(x))$, for $x \in U$, and is such that

$F(x)=C(A_1(x),\ldots,A_r(x))$,

is that for each $1 \leq i \leq r$, the rank of $TBH_i(F)*(x*)$ is at most one.

In [1], Abelson states a generalization of the theorem of Leontief [15] that is the converse of Lemma 6.1. A proof of the assertion of Leontief and the of the generalization due to Abelson can be found in Appendix B.

Theorem 6.2.(Leontief and Abelson). Suppose that $F(x,y)$ is a C^{k+1}-function, $k \geq 1$, in the variables $x=(x_1,...,x_m)$ and $(y_1,...,y_n)$.

(i) A necessary condition that there exist functions $\Phi(u,v)$, $A(x)$, and $B(y)$ such that

$$F(x,y)=\Phi(A(x),B(y))$$

is that the matrices $BH(F:x;y)$ and $BH(F:y;x)$ each have rank at most one.

(ii) If for some $1 \leq j \leq m$ and some $1 \leq k \leq n$, and some point $(x_0,y_0) \in X \times Y$

$D(x_j;F(x,y_0)) \neq 0$

and

$D(y_k;F(x_0,y)) \neq 0$,

then the matrix rank conditions of (i) are also sufficient for the existence of C^k-functions Φ, A, and B satisfying the relation $F=\Phi(A,B)$ in a neighborhood of (x_0,y_0).

Section II

Differentiable Separability

Lemma 4.1 can be used to characterize a special type of mechanism in which the message spaces are products. The most elementary form of a mechanism in which each agent has his own message space is one in which each agent reveals his parameters. A mechanism of this kind allows for the possibility that not all the individuals parameters are revealed. Because these mechanisms have message spaces that are not of minimum dimension, they are not interesting for the study of communication. They do play a significant role in establishing lower bounds for computation time. We give the following definition.

<u>Definition 6.1.</u> Suppose that X_i, $1 \leq i \leq n$, and Z are sets and suppose that $F: X_1 x \ldots x X_n \text{--->} Z$ is a function. An <u>adequate revelation mechanism realizing F</u> is a triple $(g_1 x \ldots x g_n, M_1 x \ldots x M_n, h)$ that consists of:

(i) a product of sets $M_1 x \ldots x M_n$,

(ii) a collection of functions $g_i: X_i \text{--->} M_i$, $1 \leq i \leq n$,

(iii) a function $h: M_1 x \ldots x M_n \text{--->} Z$, such that for each $(y_1, \ldots, y_n) \in X_1 x \ldots x X_n$, $F(y_1, \ldots, y_n) = h(g_1(y_1), \ldots, g_n(y_n))$.

Using the notation of Chapter IV, Lemma 4.1, the triple $(q_1 x \ldots x q_n, (X_1/F) x \ldots x (X_n/F), F^*)$ is an adequate revelation mechanism called the <u>essential revelation mechanism</u>.

In case that $(g_1 x \ldots x g_n, M_1 x \ldots x M_n, h)$ is an adequate revelation mechanism, then $M_1 x \ldots x M_n$ is an <u>adequate revelation message space</u>. The map $g_1 x \ldots x g_n$

is the message function of the adequate revelation mechanism.

The following theorem is a restatement of Lemma 4.1 in terms of adequate revelation mechanisms. It establishes the sense in which the essential revelation mechanism is the smallest adequate revelation mechanism.

Theorem 6.3. Suppose that X_i, $1 \le i \le n$, and Z are nonempty sets and suppose that $F: X_1 x \ldots x X_n \text{----}> Z$ is a function.

(i) The triple

$(q_1 x \ldots x q_n, (X_1/F) x \ldots x (X_n/F), F^*)$

is an adequate revelation mechanism that realizes F.

(ii) The message function for any other adequate revelation mechanism factors through $(X_1/F) x \ldots x (X_n/F)$.

(iii) The set $(X_1/F) x \ldots x (X_n/F)$ is the smallest set in cardinality that can be used as an adequate revelation message space for a mechanism that realizes F.

(iv) Finally, the essential revelation mechanism is the unique adequate revelation mechanism through which factor all adequate revelation mechanisms that realize F.

Section III

As we remarked in the introduction to this chapter, when the sets (X_i/F) are Hausdorff there are conditions that make (X_i/F) into topological manifolds, i.e. C^0-manifolds. In general (X_i/F) is not such a manifold. When (X_i/F) is a topological manifold, the essential revelation mechanism can be used to establish a lower bound for computation time. In this section we introduce the concept of differentiable separability and explore some of its consequences. When differentiable separability can be established it is possible to place simple global conditions on a function F to ensure that the essential revelation mechanism can be given a topological structure in which the sets (X_i/F) are topological manifolds. In order that (X_i/F) have the appropriate topological structure we start with a function defined on a differentiable manifold. Therefore, we give some concepts from differential geometry (c.f.[7]).

Definition 6.2. Let X and Y be differentiable manifolds. Let $\Phi: X \text{---}> Y$ be a differentiable mapping. If at a point $p \in X$ the mapping Φ has maximum rank, and if $\dim X \geq \dim Y$, then Φ is said to be a submersion at p. If Φ is a submersion at each point of X, then Φ is a submersion.

If a map $g:X \dashrightarrow Y$ is a submersion, then it is known(c.f. [7, p.9]) that the map can be linearized (rectified). That is, if dim(X)=n, dim Y=m, and if $p \in X$, we can choose coordinates $x_1,\ldots,x_n$ at p in a neighborhood U of p, and coordinates $y_1,\ldots,y_m$, in a neighborhood of g(p) so that for each $q \in U$, $g(q)=(x_1(q),\ldots,x_m(q))$.

<u>Definition 6.3.</u> Suppose that $X_1,\ldots,X_n$ are differentiable manifolds, where for each $1 \le i \le n$, X_i has dimension d(i). Suppose that $p_i \in X_i$, $1 \le i \le n$ and suppose that for each i,

$$\varphi_{i\ 1},\ldots,\varphi_{i\ d(i)}$$

is a coordinate system in an open neighborhood U_i of p_i. Suppose that $F:\prod_1^n X_i \dashrightarrow R$ is a C^2-function. Assume that for $1 \le i \le n$, $\varphi_i=\prod \varphi_{i\ j}$ maps U_i into an open neighborhood V_i of the origin 0_i of a Euclidean space $E^i=R^{d(i)}$ and that φ_i carries p_i to 0_i. We assume that E^i has coordinates $x_{i\ 1},\ldots,x_{i\ d(i)}$. The function F is said to be <u>differentiably separable of rank $(r_1,\ldots,r_n)$ at the point $(p_1,\ldots,p_n)$ in the coordinate system $\varphi_{1\ 1},\ldots,\varphi_{n\ d(n)}$</u> if for each $1 \le i \le n$, the matrices

$$BH(F\cdot(\prod \varphi_t)^{-1}:x_{i\ 1},\ldots,x_{i\ d(i)};x_{<-i>})$$

and

$$H^*(F\cdot(\prod \varphi_t)^{-1}:x_{i\ 1},\ldots,x_{i\ d(i)};x_{<-i>})[x_i,0_{<-i>}]$$

have rank r_i in a neighborhood of $(0_1,\ldots,0_n)$. If F is differentiably separable of rank $(r_1,\dots,r_n)$ at $(p_1,\dots,p_n)$, and if $r_i = \dim X_i$ for each $1 \le i \le n$, then we will say that F is <u>differentiably separable at $(p_1,\dots,p_n)$</u>.

The following lemma notes that the ranks of the Hessians used in the previous definition are unchanged by coordinate changes. The proof is a simple computation.

<u>Lemma 6.2.</u> Suppose that for $1 \le i \le n$, X_i and Y_i are C^2-manifolds and suppose that $h_i : Y_i \longrightarrow X_i$ is a C^2-diffeomorphism. Assume that $g : \prod_1^n Y_i \longrightarrow R$ and $F : \prod_1^n X_i \longrightarrow R$ are C^2-functions such that $g = \prod h_i \cdot F$. Suppose that $(q_1,\dots,q_n) \in \prod Y_i$ and let $h_i(q_i) = (p_i)$. If F is differentiably separable of rank $(r_1,\dots,r_n)$ at $(p_1,\dots,p_n)$, then g is differentiably separable of rank $(r_1,\dots,r_n)$ at $(q_1,\dots,q_n)$.

We can now define the term differentiably separable for a function defined on a differentiable manifold.

<u>Definition 6.4.</u> If X_i, $1 \le i \le n$, are C^2-manifolds, the function $F : X_1 \times \dots \times X_n \longrightarrow R$ is <u>differentiably separable of rank $(r_1,\dots,r_n)$ at the point $(p_1,\dots,p_n)$</u> if there is a coordinate system $\{\varphi_{i\ j}\}$ at the point $(p_1,\dots,p_n)$

such that F is differentiably separable of rank $(r_1,\ldots,r_n)$ at the point $(p_1,\ldots,p_n)$ in the coordinate system $\varphi_{1\,1},\ldots,\varphi_{n\,d(n)}$.

If $F:X_1x\ldots xX_n$--->R is differentiably separable of rank $(r(1),\ldots,r(n))$ at a point $(p_1,\ldots,p_n)$, then it is possible to write F as a function of variables $\{y_{1\,1},\ldots,y_{1\,r(1)},\ldots y_{n\,1},\ldots,y_{n\,r(n)}\}$. This assertion, Lemma 6.3, is a restatement of Theorem B.4. The proof of Theorem B.4 can be found in Appendix B together with an example of the construction.

<u>Lemma 6.3</u>. Suppose that for $1\leq i\leq n$, X_i is a C^{k+1}-manifold, $k\geq 2$. Assume,

(i) $F:X_1x\ldots xX_n$--->R is a C^{k+1}-function,

(ii) $(p_1,\ldots,p_n)$ is a point on $X_1x\ldots xX_n$.

A necessary condition that F can be written in the form

$G(\ y_{1\,1},\ldots,\cdot y_{1\,r(1)},\ldots,y_{n\,1},\ldots,y_{n\,r(n)}\ )$,

where $\{y_{i\,1},\ldots,y_{i\,d(i)}\}$ is a coordinate system on X_i, is that F is differentiably separable at $(p_1,\ldots,p_n)$ of rank $(s(1),\ldots,s(n))$ where for each $1\leq j\leq n$, $s(j)\leq r(j)$.

Conditions (i) and (ii) are also sufficient for F to be written in the form

$G(\ y_{1\,1},\ldots,\cdot y_{1\,r(1)},\ldots,y_{n\,1},\ldots,y_{n\,r(n)}\ )$,

for a C^k-function G in a neighborhood of a point

$(p_1,...,p_n)$, if F is differentiably separable of rank exactly $(r(1),...,r(n))$ at $(p_1,...,p_n)$.

Lemma 6.3 suggests that in the case of a differentiable function F satisfying the rank conditions stated in the lemma, it is possible to construct an essential revelation mechanism whose message space is a topological manifold. We now carry out the construction suggested by the lemma. The main result is given in Theorem 6.5 and in Corollary 6.5.1.

<u>Definition 6.5</u>. Suppose that X_i, $1 \le i \le n$ and Z are C^k-manifolds and suppose that $F: X_1 x \ldots x X_n \text{--->} Z$ is a differentiable function. The triple

$(g_1,...,g_n, M_1 x \ldots x M_n, h)$

that consists of spaces $M_1 x \ldots x M_n$, maps $g_1,...,g_n$, $g_i: X_i \text{--->} M_i$, $1 \le i \le n$, and function $h: M_1 x \ldots x M_n \text{--->} Z$ is an <u>adequate C^k-revelation mechanism that realizes F</u> if;

(i) each of the spaces M_i is a C^k-manifold,

(ii) each of the functions g_i, $1 \le i \le n$, and h is a C^k-differentiable function,

(iii) each g_i, $1 \le i \le n$, has a local thread at each point of M_i.

<u>Definition 6.6.</u> Suppose that $F: X_1 x \ldots x X_n \text{--->} Z$ is a differentiable map from a product of differentiable

manifolds $X_1, \ldots, X_n$ to a differentiable manifold Y. The function <u>F factors through a product of manifolds</u> $Z_1 \times \ldots \times Z_n$ if there are submersions $g_i: X_i \longrightarrow Z_i$, and a differentiable mapping $h: Z_1 \times \ldots \times Z_n \longrightarrow Y$ such that the diagram in Figure 6.1 commutes.

$$
\begin{array}{ccccc}
X_1 & \times \ldots \times & X_n & \overset{F}{\longrightarrow} & Y \\
\downarrow g_1 & & \downarrow g_n & \nearrow h & \\
Z_1 & \times \ldots \times & Z_n & &
\end{array}
$$

<u>Figure 6.1.</u>

It has not been established that the essential revelation mechanism is an adequate C^k-revelation mechanism, because the construction given in Theorem 6.3 ignores all topological and differentiable structure. The topological structure required on the spaces (X_i/F) is inherited from separator sets for F.

We begin the discussion of the topological properties of essential revelation mechanisms by studying separator sets in a special case.[6)]

If $F:X_1 x \ldots x X_n \dashrightarrow R$ is a differentiably separable function, then the function F has X_i itself as a separator set in X_i. The proof follows as a corollary to the following result. In this theorem, and the corollary that follows, the function F is assumed to be differentiably separable at every point in an open set $U_1 x \ldots x U_n$ in $X_1 x \ldots x X_n$.

<u>Theorem 6.4</u>. Suppose that X_i, $1 \le i \le n$, is a Euclidean space of dimension $d(i) \ge 1$. Suppose that for each $1 \le i \le N$, U_i is an open neighborhood of the origin 0_i of X_i and suppose that F is a C^3-function differentiably separable at each point $(p_1, \ldots, p_n) \in U_1 x \ldots x U_n$. There is an open neighborhood U of p_i such that for each pair of points x and x′ in U, $x \neq x'$, then there is a point $w \in U_{<-i>}$ such that $F(x,w) \neq F(x',w)$.

Proof. The matrix $H(F:x,y)[0,0]$ has rank d(i), by assumption. Set $X=X_i$, set $X_{<-i>}=Y$, set dim($X_{<-}$

6) In the case that F is a function from a product $X_1 x \ldots x X_n$ to a manifold Y, then the study of essential revelation mechanisms requires a more elaborate notation and a slightly more general version of Lemma 6.3. The more general version of Lemma 6.3 is given in Theorem B.4.

$_{i>}$)=N, and set m=d(i). We can change coordinates in X and Y separately to coordinates z in X and w in Y so that the new matrix H(F:z;w)[0,0] has a 1 in the z_j x w_j position, $1 \le j \le m$, and zero in all the other positions.

The Taylor series expansion for

$F(z_1,...,z_m,w_1,...,w_N)$ then has the form

$$F(z,w)=$$

$$F(0,0)+u\cdot z+v'\cdot w+w\cdot z+z^TQz+w^TQ'w+P(z^*,w^*)[z,w]$$

where Q and Q' are square matrices, u and v' are vectors in R^m and R^N respectively, $v'\cdot w$ denotes inner product, z^T denotes the transpose of the column vector z, and where P(z*,w*)[z,w] is a cubic polynomial in the variables $(z_1,...,z_m,w_1,...,w_N)$ with coefficients that are continuous functions on U x V evaluated at some point $z^* \in U$ and $w^* \in V$. These coefficients are bounded on a ball that is a compact neighborhood of $(0,0) \in U' \times V'$, $U' \subseteq U$ and $V' \subseteq V$. Then for $z,z' \in U'$ and $w \in V'$,

$$|F(z,w)-F(z',w)|=$$

$$|u.(z-z')+w.(z-z')+z^TQz-z'^TQz' +$$

$$P(z'^*,w'^*)[z',w]+P(z^*,w^*)[z,w]|.$$

The vector $(z-z') \neq 0$ and the w is to be chosen in the set V'. Set $z'^TQz'-z^TQz=K$, set $u\cdot v=L$, and set $(z-z')=v$. To complete the proof, it will suffice to show that the function

$$w \cdot v+P(z'^*,w'^*)[z',w]+P(z^*,w^*)[z,w]+K+L$$

is not constant on the ball V'. For this it will suffice to show that the function

$$Q=w \cdot v+P(z'^*,w'^*)[z',w]+P(z^*,w^*)[z,w]$$

is not constant on the ball V'. The function

$$P(z'^*,w'^*)[z',w]+P(z^*,w^*)[z,w]$$

is a homogeneous cubic $\Sigma\, a_{\alpha\,\beta}\, z^{\alpha}\, w^{\beta}$ in the variables $w_1,\ldots,w_N$ with coefficients

$$\{a_{\alpha\,\beta}(z,z',w,w')\}$$

that are functions bounded on U' x V'.

Set w=tv. The powers of the constants $z_1,\ldots,z_m$ can be combined with the coefficients $a_{\alpha\,\beta}$ and therefore

$$Q=t|v|^2+a(t)\,t^3,$$

where the a(t) is also bounded as a function of t.

If a(t)=0 identically in t, then because v≠0, different values of t produce different values of Q.

If a(t)≠0, and

$$|v|^2+a(t)t^2=c \text{ (a constant)},$$

then

$$a(t)=(c-|v|^2)/t^2,$$

and therefore a(t) is not bounded as t approaches 0. Therefore Q is not a constant.▒

We now give conditions on a function F that is differentiably separable of rank $(r_1,\ldots,r_n)$, so that the sets (X_i/F), with the quotient topology, have the structure of a C^0-manifold of dimension r_i. Under these conditions the set theoretic essential revelation mechanism is a topological essential revelation mechanism.

Definition 6.7. If X_i, $1 \le i \le n$, are topological spaces, then a real valued function

$F: X_1 x \ldots x X_n \dashrightarrow R$

induces strong equivalence on X_i, if the following condition is satisfied for each $x, x' \in X_i$, such that $x \neq x'$;

(i) if there is an open neighborhood U of a point $q \in X_{<-i>}$, such that $F(x \int_i u) = F(x' \int_i u)$ for each $u \in U$, then $F(x \int_i z) = F(x' \int_i z)$ for all $z \in X_{<-i>}$.

It is relatively easy to find classes of functions that induce strong equivalence. Suppose the X_i are Euclidean spaces with coordinates $x_{i\ j}$, $1 \le i \le n$, $1 \le j \le d(i)$. If for each $1 \le i \le n$, $\beta(i) = (\beta(i\ 1), \ldots, \beta(i\ d(i)))$ is a sequence of nonnegative integers, denote by $x_i^{\beta(i)}$ the monomial

$$x_{i\ 1}^{\beta(i\ 1)} \ldots x_{i\ d(i)}^{\beta(i\ d(i))},$$

and denote by

$$x_1^{\beta(1)}\ldots x_n^{\beta(n)}$$

the product of the monomials $x_i^{\beta(i)}$. Write

$$F(x_1,\ldots,x_n)=$$

$$\Sigma_{\beta(1),\ldots,\beta(n)}\, A_{\beta(1)\ldots\beta(n)}(x_1)\, x_2^{\beta(2)}\ldots x_n^{\beta(n)},$$

where the $A_\beta(x_1)$ are polynomials in x_1. Then for x_1, x'_1 in X_1,

$$F(x_1,x_{<-1>})=F(x'_1,x_{<-1>})$$

for $x_{<-1>}$ in an open set in $X_{<-1>}$, if and only if

$$\Sigma\,[A_\beta(x_1)-A_\beta(x'_1)]x_2^{\beta(2)}\ldots x_n^{\beta(n)}=0$$

for the $x_2,\ldots,x_n$ chosen arbitrarily in an open set in $X_2x\ldots xX_n$. However, a polynomial vanishes in an open set if and only if each of its coefficients is zero. Therefore if

$$F(x_1,x_{<-1>})=F(x'_1,x_{<-1>})$$

for the $x_{<-1>}$ chosen in some open set, it follows that for each β,

$$A_\beta(x_1)-A_\beta(x'_1)=0.$$

That is, F induces a strong equivalence relation on X_1.

<u>Theorem 6.5.</u> Suppose that X_i, $1\le i\le n$ are C^4 manifolds of dimensions $d(1),\ldots,d(n)$, respectively. Suppose that $F:X_1x\ldots xX_n$---->R is a C^4 function that is differentiably separable on $X_1x\ldots xX_n$ of rank $(r(1),\ldots,r(n))$ where each $r_i\ge 1$. Assume that F induces strong equivalence in X_i for each i. If

(i) the spaces (X_i/F) are all Hausdorff,

(ii) quotient map $q_i:X_i \dashrightarrow (X_i/F)$ is open for each $1\le i\le n$,

then, for each $1\le i\le n$, the space (X_i/F) (with quotient topology) is a topological manifold (i.e. a C^0-manifold). Furthermore, the quotient map

$$q_i:X_i \dashrightarrow (X_i/F)$$

has a local thread in the neighborhood of each point.

Proof. Suppose that $p_i{*}\in(X_i/F)$, $1\le i\le n$. Choose a point $p_i\in X_i$, $1\le i\le n$, such that $q_i(\ p_i\ )=p_i{*}$. Because the function F is differentiably separable of rank $(r(1),\ldots,r(n))$ at the point $(p_1,\ldots,p_n)$, it follows from Lemma 6.3 that for $1\le i\le n$, there is an open neighborhood U_i of $p_{<-i>}$ in $X_{<-i>}$, a coordinate system $x_i=(x_{i\ 1},\ldots,x_{i\ d(i)})$ in X_i such that $x_i(\ p_i\ )=(0,\ldots,0)$ and a C^3-function G defined in a neighborhood of the origin, such that

$$F(\ x_1,\ldots,x_n\ )=G(\ (x_{i\ 1},\ldots,x_{i\ r(i)})\textstyle\int_i z\ )$$

for each $z\in U_{<-i>}$.

Denote by $S{*}_i$ the set of elements $(x_{i\ 1},\ldots,x_{i\ r(i)},0,\ldots,0)$ that lie in U_i. Choose in $S{*}_i$ a compact neighborhood S_i of $(0,\ldots,0)$ (in the induced topology on $S{*}_i$.) The map q_i carries the set U_i to an open set of (X_i/F) because we have assumed that q_i is an open map. We have assumed that the equivalence relation induced on $X_{<-i>}$ by F is strong, therefore the equality

$F(x_{i\,1},\dots,x_{i\,r(i)},b_1,\dots,b_{d(i)-r(i)} \int_i z_{<-i>})=$

$F((x_{i\,1},\dots,x_{i\,r(i)},0,\dots,0) \int_i z_{<-i>})$

implies that

$q_i(x_{i\,1},\dots,x_{i\,d(i)})=q_i(x_{i\,1},\dots,x_{i\,r(i)})$

for each $(x_{i\,1},\dots,x_{i\,d(i)})$ in U_i. Therefore,

$q_i(U_i)=q_i(S^*_i)$.

The set S^*_i was constructed so that q_i is one-to-one on S^*_i. By assumption, the space (X_i/F) is Hausdorff, therefore the restriction of q_i to S_i is a homeomorphism from S_i to a neighborhood N_i of p^*_i. Denote by s_i the inverse of q_i in N_i. It follows that the point $p^*_i \in X_i$ has a neighborhood N_i that is homeomorphic to a neighborhood of the origin of the space $R^{r(i)}$. Furthermore, the function s_i is a thread of q_i on the set N_i.▒

The following corollary states that the essential revelation mechanism is a C^0-essential revelation mechanism. In this case, under the assumptions placed on F, each C^0-adequate revelation mechanism factors through the C^0-essential revelation mechanism.

<u>Corollary 6.5.1</u>. Suppose that X_i, $1 \le i \le n$ are C^4-manifolds and that X_i has dimension $d(i)$. Assume that $F:X_1 x \dots x X_n \text{----}> R$ is a real valued function on F that satisfies the following conditions:

(i) there are integers $(r(1),\dots,r(n))$,

$1 \le r(i) \le d(i)$, such that at each point $(p_1,\ldots,p_n) \in X_1 x \ldots x X_n$, F is differentiably separable of rank $(r(1),\ldots,r(n))$,

(ii) for each i, the map $q_i: X_i \text{--->} (X_i/F)$ is open and (X_i/F) is Hausdorff,

(iii) for each i, F induces a strong equivalence relation on X_i.

Then the triple

$(q_1 x \ldots x q_n, (X_1/F) x \ldots x (X_n/F), F^*)$

where;

(1) each (X_i/F) is given the quotient topology,

(2) the maps $q_i: X_i \text{--->} (X_i/F)$ is the quotient map,

(3) $F^*: (X_1/F) x \ldots x (X_n/F) \text{--->} R$ is the function such that

$F^*(q_1(x_1),\ldots,q_n(x_n))=$

$F(x_1,\ldots,x_n)$

for each $(x_1,\ldots,x_n) \in X_1 x \ldots x X_n$,

is an adequate C^0-revelation mechanism that realizes F. The space (X_i/F) has dimension r(i). Furthermore, if a triple

$(g_1 x \ldots x g_n, Z_1 x \ldots x Z_n, G)$

is such that

$g_i: X_i \text{--->} Z_i$,

$G: Z_1 x \ldots x Z_n \text{--->} R$,

and the triple is an adequate revelation mechanism that realizes F, then there are continuous maps

$g*_i:Z_i\text{--->}(X_i/F)$

such that the diagram in Figure 6.2 commutes.

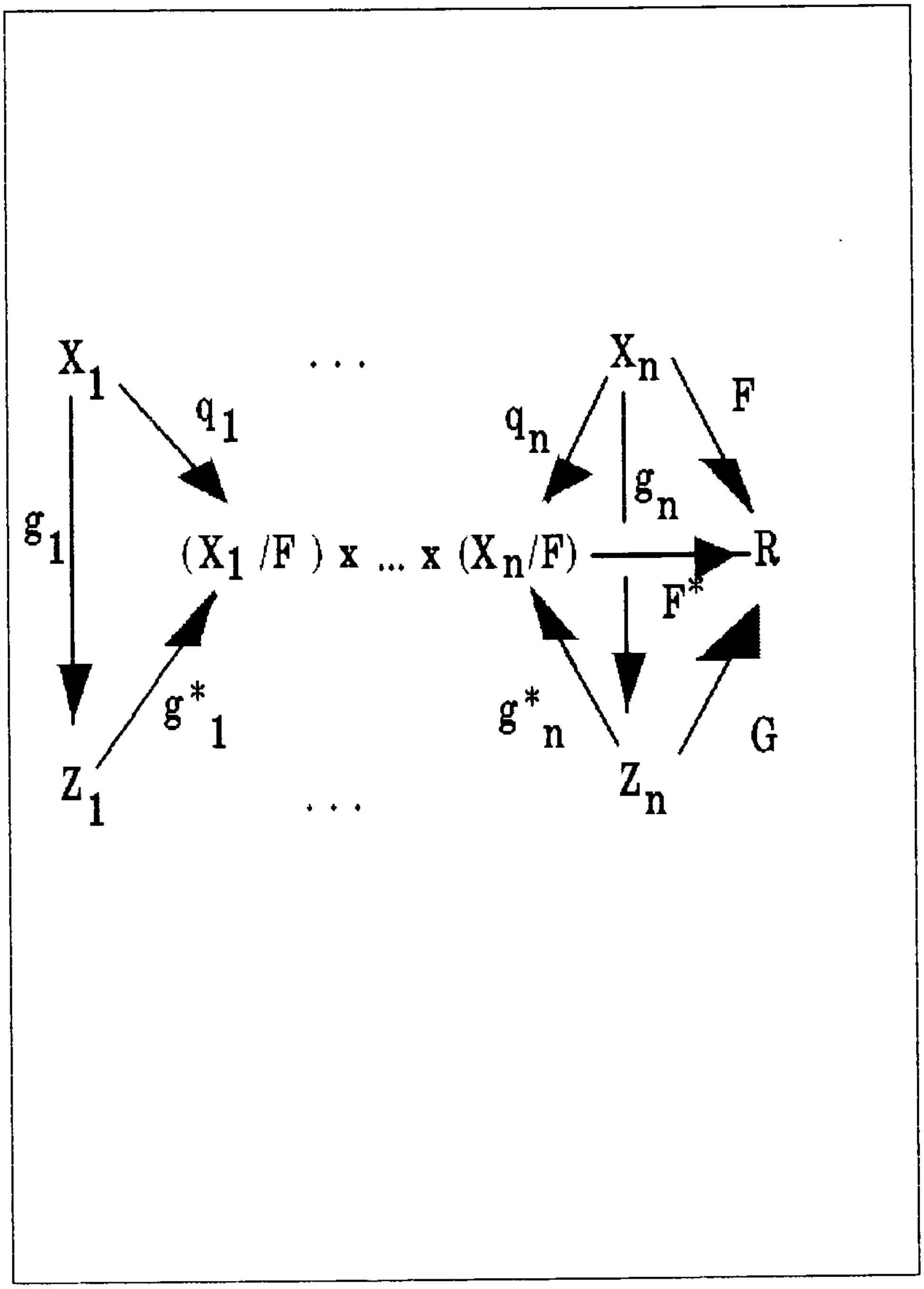

Figure 6.2

Proof. We have already shown in Theorem 6.5 that the triple $(q_1 x \ldots x q_n, (X_1/F) x \ldots x (X_n/F), F^*)$, is an adequate revelation mechanism that realizes F. Suppose that $z^*_i \in Z_i$. Denote

$$(g_1(w), \ldots, g_{i-1}(w), g_{i+1}(w), \ldots, g_n(w))$$

by $g_{<-i>}(w)$, for each $w \in X_{<-i>}$. Choose an element $x^*_i \in X_i$ such that

$$g_i(x^*_i)=z^*_i.$$

Suppose that x'_i, $x^*_i \in X_i$, such that

$$g_i(x^*_i)=g_i(x'_i)=z^*_i.$$

Then for each $w \in X_{<-i>}$,

$$F(x^*_i \int_i w)=G(g_i(x^*_i) \int_i g_{<-i>}(w))=$$

$$G(g_i(x'_i) \int_i g_{<-i>}(w))=$$

$$F(x'_i \int_i w).$$

Therefore

$$q_i(x^*_i)=q_i(x'_i).$$

Set

$$g^*_i(z^*_i)=q_i(x^*_i).$$

Because the map $g_i:X_i \text{--->} Z_i$ has a thread in the neighborhood of each point, there is a neighborhood N of the point z^*_i and a thread $s_i:N \text{--->} X_i$ such that

$$g_i(s_i(z^*))=g_i(z^*)$$

for each $z^* \in N$.

Then

$$g^*_i(z^*)=q_i(s_i(z^*)).$$

Because both q_i and s_i are continuous, it follo... that the map g^*_i is continuous.▒

Theorem 6.5 and Corollary 6.5.1 immediately give a lower bound on the time required for the computation of a C^3-function that is differentiably separable. Each (X_i/F) in the C^0-essential revelation mechanism has dimension r_i. If $p \in (X_i/F)$ there is a local thread, s, from (X_i/F) into X_i defined on a neighborhood of p. If U_i is a compact neighborhood of p on which s is defined, then the image of U_i under s is a locally Euclidean subspace of X_i that is a separator set of F in X_i. The image of U_i has the same dimension as U_i. Therefore, the minimum time required to compute an encoded version of F is, by the Dimension Based Lower Bound Theorem, $\Sigma\, r_i$.

When F satisfies the conditions given in the statement of Corollary 6.5.1, r_i is the largest dimension of locally Euclidean separator sets in X_i. Indeed, if S is a separator set in X_i, and if p is a point in S, then the quotient map q_i carries S into X_i. Suppose that $q_i(p)=p^*$. The map q_i is one-to-one on S, because S is a separator set. Assume that U is a compact neighborhood of p. If U^* is the image under q_i of U in (X_i/F), then the subspace U^* is Hausdorff because (X_i/F) is assumed to be Hausdorff. Therefore the restriction of q_i to U is a homeomorphism. But U^*

is a subspace of a topological space of dimension r_i, therefore U* has dimension at most r_i (c.f. [10], Theorem III.1 p. 26). It is also clear that if F satisfies the conditions of Corollary 6.4.1, and if the adequate revelation mechanism

$$(g_1,\ldots,g_n,Z_1 x\ldots xZ_n,h),$$

where

$$g_i:X_i\text{--->}Z_i \quad ,1\leq i\leq n,$$

and

$$h:Z_1 x\ldots xZ_n\text{--->}R,$$

realizes F, then the minimum computing time required for h is at least $\Sigma\ r_i$. This follows from the fact that the map h must factor through $(X_1/F)x\ldots x(X_n/F)$ where the factorization is given by maps

$$h_i:Z_i\text{--->}(X_i/F)$$

(c.f. Figure 6.2) and the fact that the maps h_i are locally threaded.

The dimension of the message space of the essential revelation mechanism is also an upper bound on the dimension of the minimal message space, but it is not as good as the bound given by parameter transfer. The dimension of the essential revelation mechanism, when the essential revelation mechanism exists, is best viewed as a lower bound on computation.

Chapter VII

Computational Complexity of an Edgeworth Box Economy with a Walrasian Performance Standard

In this chapter we study the efficient frontier, introduced in Chapter V, for a particular performance function. We consider the case of two agents, each with a two dimensional parameter space (environment) with, say, coordinates (x,z) for agent 1 and (x′,z′) for agent 2. The (real-valued) performance function is given by

$$Q(x,z,x',z') = (z-z')/(x-x')$$ [7)]

---

7) The performance standard $Q(x,z,x',z') = (z-z')/(x-x')$ is a Walrasian one for the case of two agents trading two goods. Let (Y,Z) denote the holdings of the two goods. We assume utilities to be quadratic in Y and linear in Z. The initial endowments of the two goods are

$w^i_{(X)}$ and $w^i_{(Y)}$, $i=1,2$.

$$u^i(x,z) = \alpha^i Y^i + 1/2\ \beta^i (Y^i)^2 + Z^i \qquad i=1,\ 2$$

$$y^i = Y^i - w^i_{(Y)} = \text{net trade of } i^{th} \text{ agent};$$

$$y^1 + y^2 = 0.$$

Equilibrium conditions:

$$u^i = \alpha^i (y^i + w^i_{(Y)}) + 1/2\ \beta^i (y^i + w^i_{(Y)})^2 + Z^i,$$

i=1, 2,

$$\frac{du^i}{dy^i} = \alpha^i + \beta^i (y^i + w^i_{(Y)}) = p \text{ (the price)}, \qquad i=1,\ 2$$

In Section I, we ask how long it takes to compute the equilibrium message $\mu(x,z,x',z')$ of a privacy preserving mechanism realizing Q at an arbitrary parameter point (x,z,x',z') using an analytic

(2,1)-network? The question arises from the interpretation of a decentralized mechanism in

---

Let

$$y^1=y, \qquad y^2= -y$$

$$\alpha^1 + \beta^1(y + w^2{}_{(Y)})=p$$

$$\alpha^2 + \beta^2(-y + w^2{}_{(Y)})=p$$

Let

$$\gamma^i = \alpha^i + \beta^i w^i{}_{(Y)} \qquad i=1,\ 2.$$

Then the equilibrium conditions are written

$$\gamma^1 + \beta^1 y=p$$

$$\gamma^2 - \beta^2 y=p.$$

Let

$$(x,z)=(-\beta^1,\gamma^1)$$

$$(x',z')=(\beta^2,\gamma^2).$$

Then

$$z-xy=z'-x'y$$

or

$$y=(z-z')/(x-x'),$$

which is our performance standard. Hurwicz presented essentially the same derivation of this performance standard, except for changes in sign, in [11].

equilibrium form as a one step iterative process in which the outcome function is assumed to be a projection (c.f. Assumption 5.2 in Chapter 5 and the discussion that precedes it.) In Section I, no coordinate changes are allowed either in the message space or in the agents' parameter spaces. This restriction makes the analysis of the computation particularly easy. In Section III, we use the results of Section I to analyze the efficient frontier for the function Q.

In Section II of this chapter we suppose that each agent may independently make a real linear transformation of his parameter space corresponding to different encodings of his parameters. As we have seen in Section III of Chapter IV, computation time may well depend on the particular coordinate systems used in each of the three spaces involved in the problem, namely, the parameter spaces of the agents and the message space. But these coordinate systems are not necessarily intrinsic. In the case of the message space, the designer of the mechanism is free to specify the coordinate system. In the case of the parameter space of an agent, the agent's perception or experience of his environment e.g., his preferences, is presumably what is intrinsic. The particular choice of coordinates is an artifact of modelling. Therefore, we introduce into the problem the possibility of different coordinate systems separately in each space. In this chapter we also consider coordinate changes in the message space that are linear. In Chapter VIII we consider coordinate changes in the agents' parameter spaces and in the message space that are real analytic transformations.

Section I.

Complexity of Computing the Walrasian Equilibrium

In this section we study the mechanism that has as its message correspondence

$$\mu(x,z,x',z')=\left(\frac{z-z'}{x-x'}, \frac{xz'-x'z}{x-x'} \right).$$

It is clear that the computation of μ can be done in three units of time by analytic (2.1)-networks. Namely, the network shown in Figure 7.1 computes Q in 2 units of time, while the network shown in Figure 7.2 computes P in 3 units of time.

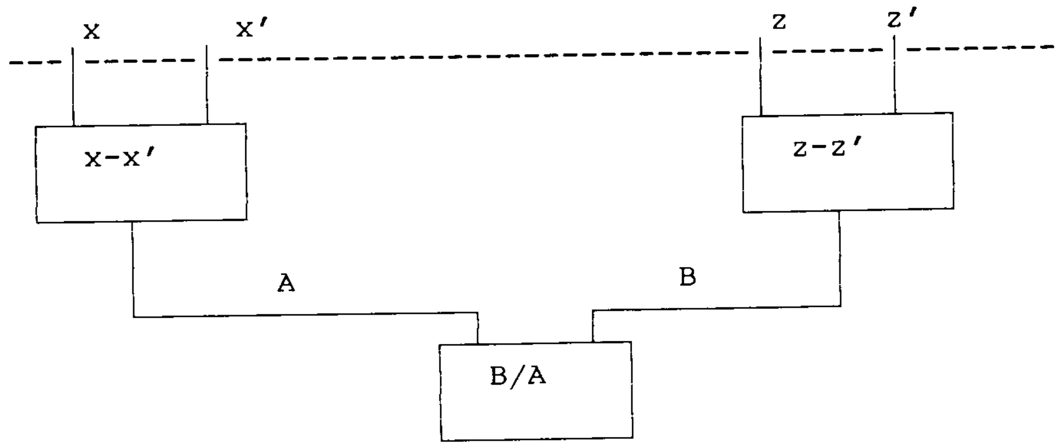

Figure 7.1

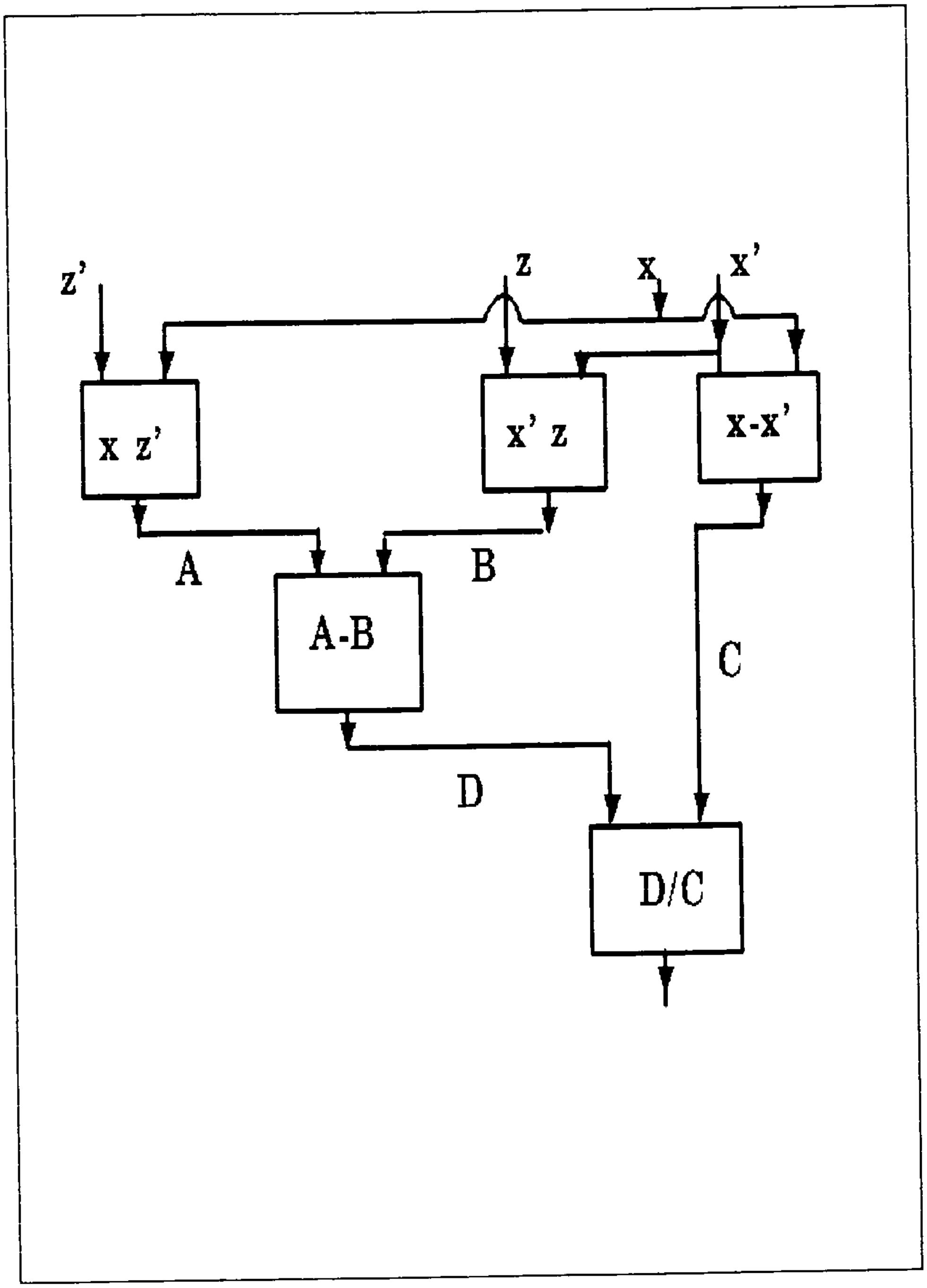

Figure 7.2

So, the question of the time required to compute μ is reduced to whether there is a (2,1)-network N that computes μ in two units of time, allowing for linear coordinate transformations of the message space and the two parameter spaces. Such a network N would be of the form displayed in Figure 7.3, where A, B, C, D, E, and F are real analytic functions and x, y, z, and w are real variables.

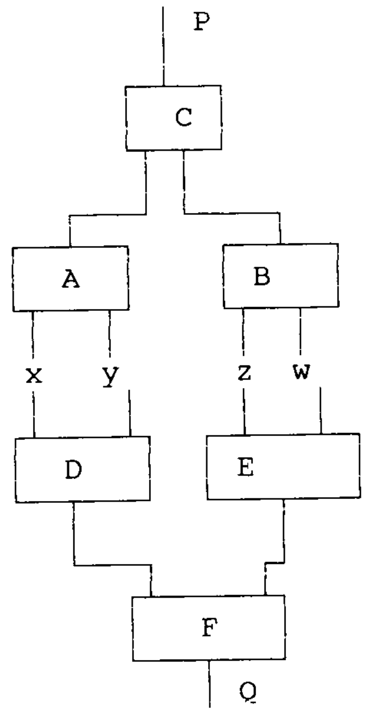

Figure 7.3

Theorem 6.1 in Chapter VI states necessary conditions that an analytic function $F(x,y,z,w)$ can be written in the form

$$F(x,y,z,w)=C(A(x,y),B(z,w)).$$

We use those conditions to prove that no (2,1)-network with real analytic modules can compute both components of μ in two units of time.

<u>Notation</u>. If $T=(f_{ij}(x_1,\ldots,x_n))$ is a matrix of functions of the real variables $(x_1,\ldots,x_n)$, and if $a=(a_1,\ldots,a_n)$ is an n-tuple of real numbers, then $T(a)$ denotes the matrix with entries $(f_{ij}(a))$.

Theorem 7.1 states that the time required to compute an encoded version of the message $\mu(x,z,x',z')$ is at least 3 units of time. Definition 4.3 of Chapter IV, defines the concept of a network computing an encoded version of a function $F:X_1 x\ldots xX_n \text{--->} Y$. In Theorem 7.1, $X_1=X_2=R^2$, and $Y=M=R^2$. We suppose that the encoding functions for the network are

$$g^i:R^2\text{--->}R^2,\ i=1,\ 2,$$

where each g^i is the identity function. We suppose that M is encoded by functions

$$(k_1,k_2):R^2\text{--->}R^2,$$

where

$$k_1(m^1,m^2)=m^1$$

and

$k_2(m^1,m^2)=m^2$.

Theorem 7.1. Suppose that X_1 and X_2 are Euclidean spaces of dimension 2 with coordinates (x,z) and (x',z'), respectively. Suppose that Q is the performance function

$$Q(x,z,x',z')=\frac{(z-z')}{(x-x')}.$$

Suppose that

$$P(x,z,x',z')=\frac{(xz'-x'z)}{(x-x')},$$

suppose that M is the Euclidean space R^2 with coordinates m^1, m^2, and suppose that h is the projection[8)]

$$h(m^1,m^2)=m^1.$$

Assume that Q is realized by a mechanism (μ,M,h) where

$$\mu(x,z,x',z')=(Q,P).$$

If N is an analytic (2,1)-network that computes an encoded version of μ, where the encodings $g^1:X_1\text{--->}R^2$ and $g^2:X_2\text{--->}R^2$ are the identity functions, then network N requires 3 units of time for the computation.

Proof. A coordinate change in the X_i that is a translation does not effect computing time. Indeed, suppose that the original network computes a function in time t, and that the computation is represented by a directed graph that is a tree. Suppose a pair of input vertices have associated variables r and s, and these

8) See Assumption 5.2, Chapter V.

variables are connected by edges e_1 and e_2, respectively, to a module $g(e_1,e_2)$. If the variables r and s are translated to $r'=r+a$ and $s'=s+b$, then we construct a new network, using the same tree as the original, and replace the module $g(e_1,e_2)$ by the module $G(e_1,e_2)=g(e_1-a,e_2-b)$. If all the modules in the tree other than those connected to input vertices are unchanged, the new network computes the same function as the original network and the new network carries out the computation in the same time as the original using the translated coordinates.

Without loss of generality we use the coordinates $R=x-1$, $T=x'+1$, $S=z$, and $U=z'$. In these coordinates

$$Q= \frac{(S-U)}{(2+R-T)}$$

and

$$P=\frac{(S+U+RU-ST)}{(2+R-T)} .$$

The network in Figure 7.1 computes Q in time 2 and the network in Figure 7.4 computes Q in time 2 using the inputs R, S, T, and U.

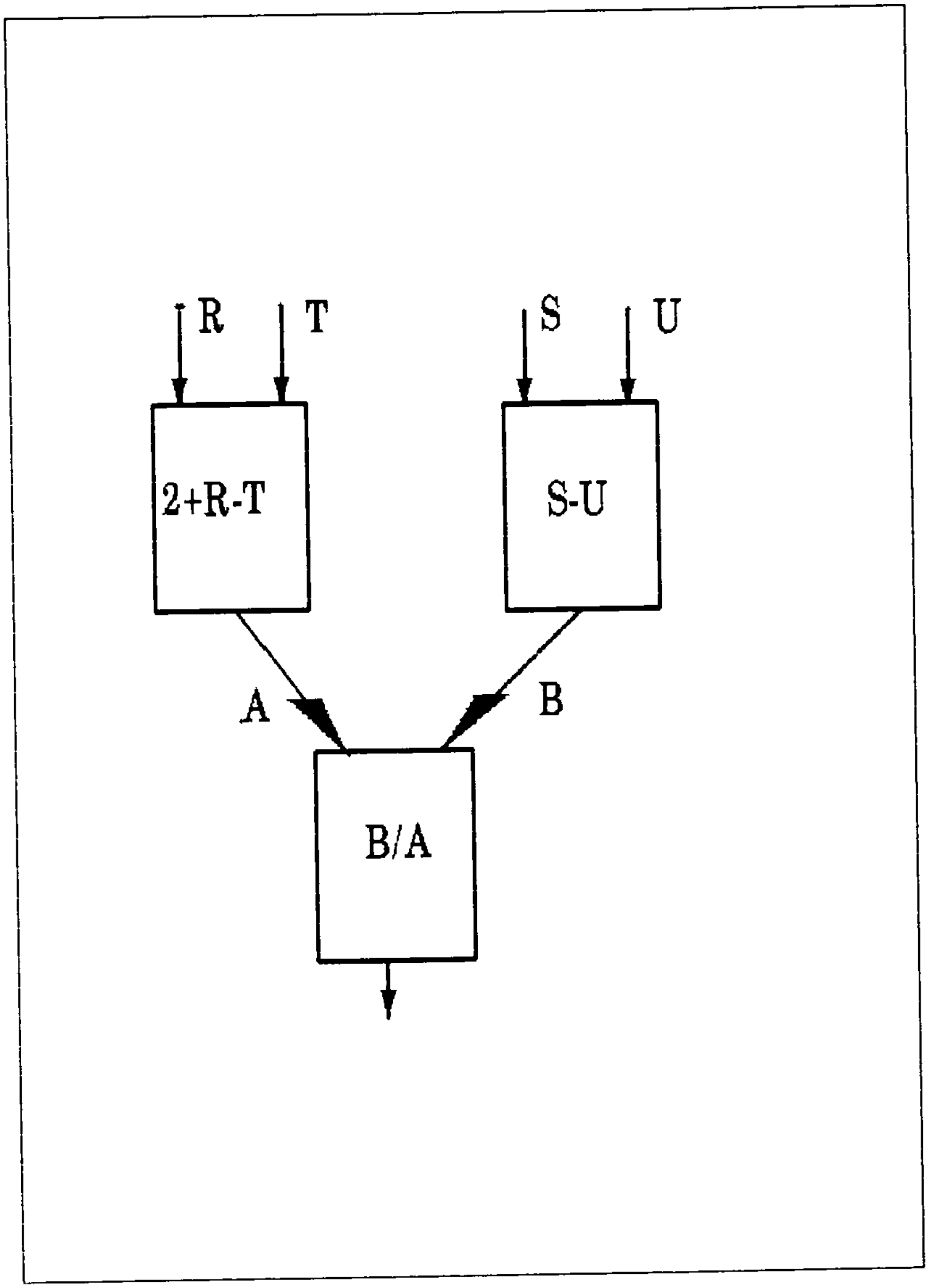

Figure 7.4

In order for a network to compute P in time 2, we must be able to write

$$P(R,S,T,U)=C'(A'(S,T),B'(R,U))$$

or

$$P(R,S,T,U)=C''(A''(R,T),B''(S,U)),$$

where A, A", B', B", C', C" are real analytic functions in the neighborhood of the origin of R^2. Theorem 6.1 states that in order for A', B', and C' to exist the matrix

$$W_1(0,0)= \left| \begin{matrix} \frac{\partial P}{\partial S} & \frac{\partial^2 P}{\partial R \partial S} & \frac{\partial^2 P}{\partial S \partial U} \\ \frac{\partial P}{\partial T} & \frac{\partial^2 P}{\partial R \partial T} & \frac{\partial^2 P}{\partial U \partial T} \end{matrix} \right|_{(0,0)}$$

must have rank at most 1.

But

$$P=(S+U+RU-TS)(\Sigma_{j=0}^{\infty}(-1)^j(1/2)^{j+1}(T-R)^j)=$$

$$(1/2)[(S+U+RU-TS)+(S+U)(T-R)/2]+ \theta,$$

where θ is a sum of monomials in R,S,T,U of degree at least 3. But then

$$W_1(0,0)= \left| \begin{matrix} 1/2 & -1/4 & 0 \\ 0 & 0 & 1/4 \end{matrix} \right|$$

has rank 2. Thus the necessary condition of Theorem 6.1 that $W_1(0,0)$ have rank at most one if P is to be computed in two units of time by an analytic (2,1)-network is not satisfied. If P can be computed in time 2 by an analytic (2,1)-network it must be the case that

$$P(R,S,T,U)=C''(A''(R,T),B''(S,U)).$$

But in this case, again by Theorem 6.1, the matrix

$$W_2(0,0)=\left|\begin{matrix} \frac{\partial P}{\partial S} & \frac{\partial^2 P}{\partial R\,\partial S} & \frac{\partial^2 P}{\partial S\,\partial T} \\ \frac{\partial P}{\partial U} & \frac{\partial^2 P}{\partial R\,\partial U} & \frac{\partial^2 P}{\partial T\,\partial U} \end{matrix}\right|(0,0)$$

can have rank at most 1. But

$$W_2(0,0)=\left|\begin{matrix} 1/2 & -1/4 & -1/4 \\ 1/2 & 1/4 & 1/4 \end{matrix}\right|$$

has rank 2. Therefore, P cannot be computed in less than 3 units of time. The network given in Figure 7.5 computes P in 3 units of time from the inputs R, S, T, U.▒

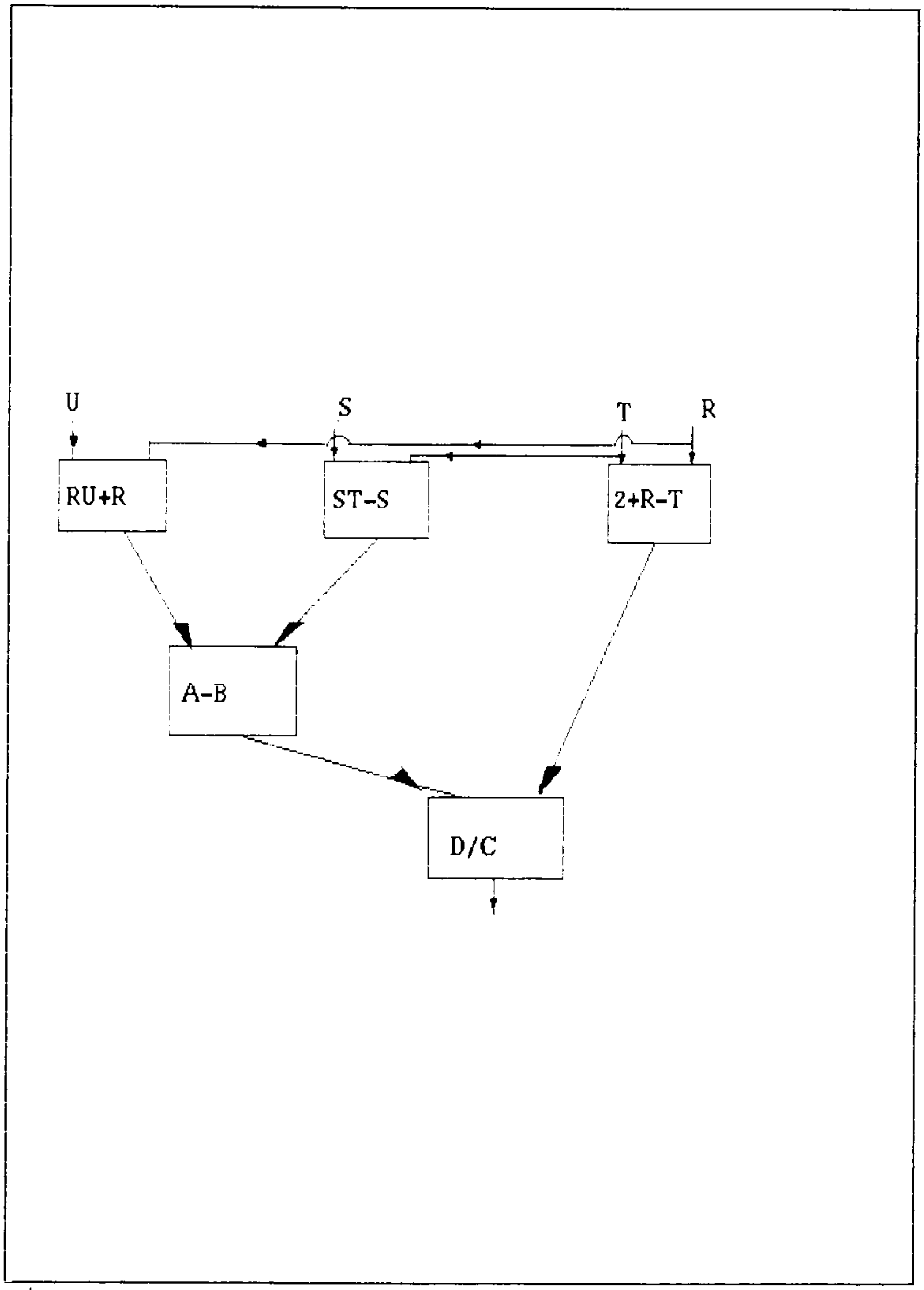

Figure 7.5

Section II.

In Section I we examined the computation of the functions P and Q by networks that receive the variables x, z, x′, and z′. In this section we first examine the task of computing the function

$$\mu(x,z,x',z')=(Q,P)$$

when linear coordinate changes are allowed in the message space $M=R^2$. We then examine the case in which the agents are also allowed to apply linear coordinate changes. As in section I, the outcome function h:M--->R is a projection.

We first show that there is no linear change of coordinates in the message space that reduces the time required for a network to compute μ.

<u>Lemma 7.1</u>. Suppose that X_1 and X_2 are Euclidean spaces of dimension 2 with coordinates (x,z) and (x′,z′), respectively. Suppose that Q is the performance function

$$Q(x,z,x',z')=(z-z')/(x-x').$$

Suppose that

$$P(x,z,x',z')=(xz'-x'z)/(x-x').$$

Suppose that M is the Euclidean space R^2 with coordinates m^1 and m^2 and assumed that Q is realized by the mechanism (μ',M,h') where

$$\mu'(x,z,x',z')=(Q,aP+bQ)$$

where $a,b \in R$ and $a \neq 0$. If N' is an analytic (2,1)-network that computes an encoded version of μ', where the encodings $g^i : X_i \text{--->} R^2$ are identity functions, then the network N' requires at least three units of time for the computation.

Proof. Return to the notation used in the proof of Theorem 7.1. In order for the network N' to compute $P'=Q+aP$ in two units of time, it must be possible to write

$$P'(R,S,T,U)=C'(A'(S,T),B'(R,U))$$

or

$$P'(R,S,T,U)=C''(A''(R,T),B''(S,U)),$$

where A',A'',B',B'',C',C'' are real analytic functions in a neighborhood of the origin of R^2. Again refer to Theorem 6.1. Set

$$Y_1 = \begin{vmatrix} \frac{\partial P'}{\partial S} & \frac{\partial^2 P'}{\partial R \partial S} & \frac{\partial^2 P'}{\partial S \partial U} \\ \frac{\partial P'}{\partial T} & \frac{\partial^2 P'}{\partial R \partial T} & \frac{\partial^2 P'}{\partial T \partial U} \end{vmatrix}$$

If $P'(R,S,T,U)=C'(A'(S,T),B'(R,U))$,

then Y_1 must have rank at most one in a neighborhood of the origin. However,

$$\left(\frac{\partial^2 P'}{\partial R \partial S}\right)\left(\frac{\partial^2 P'}{\partial T \partial U}\right)-\left(\frac{\partial^2 P'}{\partial R \partial T}\right)\left(\frac{\partial^2 P'}{\partial S \partial U}\right) =$$

$$(a+aR-b)(-a-b+aT)/(2+R-T)^2.$$

Because $a \neq 0$, this expression does not vanish identically in the neighborhood of the origin. therefore, $P' \neq C'(A'(S,T),B'(R,U))$. If P' can be computed in two units of time, then

$$P'(R,S,T,U)=C''(A''(R,T),B''(S,U)).$$

Set

$$Y_2 = \begin{vmatrix} \frac{\partial P}{\partial S} & \frac{\partial^2 P}{\partial R\, \partial S} & \frac{\partial^2 P}{\partial S\, \partial T} \\ \frac{\partial P}{\partial U} & \frac{\partial^2 P}{\partial R\, \partial U} & \frac{\partial^2 P}{\partial T\, \partial U} \end{vmatrix}$$

If $P'(R,S,T,U)=C''(A''(R,T),B''(S,U))$, then the determinant formed by the last two columns of Y_2 must be zero. However, we have already seen, in the discussion of Y_1, that the determinant in not zero. Therefore, it follows that no linear change of coordinates in the message space M can reduce the time required to compute μ to two units of time.▒

If the agents are allowed to make linear changes of coordinates in their parameter spaces, the problem is considerably more complicated. Assume that the encoding functions g^i are, as in Theorem 7.1, identity functions. Thus a network that computes P and Q must carry out the computation using the coordinates that are passed by the agents. Assume that the first agent, whose coordinates are R and S, introduces new coordinates $A_1=(r,s)$ given by the linear transformation

(A_1): $R=ar + bs$

$S=cr + ds.$

Suppose that the second agent uses new coordinates $A_2=(t,u)$ given by

(A_2): $T=et + fu$

$U=gt + hu$.

The elements a,b,c,d,e,f,g, and h are to be real numbers and the determinants

$$\text{Det} \begin{vmatrix} a & b \\ c & d \end{vmatrix} \text{ and } \text{Det} \begin{vmatrix} e & f \\ g & h \end{vmatrix}$$

are both nonzero. The following lemma uses this notation and shows that the new coordinates A_1 and A_2 cannot be chosen to decrease the time require to compute μ. The function Q plays no role in this result.

<u>Lemma 7.2</u> There is no choice of coordinates A_1 and A_2 for the parameter spaces X_1 and X_2 from which P can be computed in less that 3 units of time if the encoding functions used to compute an encoded version of P are identity functions.

Proof. If P can be computed in time 2 using as inputs the coordinate sets A_i, then either

(I) $P(R,S,T,U)=C(A(r,t),B(s,u))$

or

(II) $P(R,S,T,U)=C'(A'(r,u),B'(s,t))$.

Note also, that because the coordinate changes A_1 and A_2 are general linear changes of coordinates, if we show that no choice of a,b,c,d,e,f,g,h can be made so that

$$P(R,S,T,U)=C(A(r,t),B(s,u))$$

then no choice of a,b,c,d,e,f,g,h can be made so that

$$P(R,T,S,U)=C'(A'(r,u),B'(s,t)).$$

Theorem 6.1 gives the criterion we use to examine the possibility that (I) can to be satisfied. If (I) can be solved for the functions A, B, and C, then the matrix

$$W(r,t,s,u)= \begin{vmatrix} P_r & P_{rs} & P_{ru} \\ P_t & P_{st} & P_{tu} \end{vmatrix}$$

has rank at most one in a neighborhood of the origin, and the matrix

$$W(s,u,r,t)= \begin{vmatrix} P_s & P_{rs} & P_{st} \\ P_u & P_{rt} & P_{tu} \end{vmatrix}$$

must have rank at most 1 in the neighborhood of the origin. For the analysis of these conditions, we need the following list of derivatives.

$Q_U=-1/(2+R-T)$; $Q_S=1/(2+R-T)$;

$Q_R=-(S-U)/(2+R-T)^2$; $Q_T=(S-U)/(2+R-T)^2$;

$P_R=(T-1)(S-U)/(2+R-T)^2$; $P_S=(1-T)/(2+R-T)$;

$P_T=-(1+R)(S-U)/(2+R-T)^2$; $P_U=(1+R)/(2+R-T)$;

$Q_{RS}=-1/(2+R-T)^2$; $Q_{RU}=1/(2+R-T)^2$;

$Q_{ST}=1/(2+R-T)^2$; $Q_{TU}=-1/(2+R-T)^2$;

$Q_{SU}=0$; $Q_{RT}=-2(S-U)/(2+R-T)^3$;

$P_{RS}=(-1+T)/(2+R-T)^2$; $P_{RU}=(T-1)/(2+R-T)^2$; $P_{ST}=-(1+R)/(2+R-T)^2$; $P_{TU}=(1+R)/(2+R-T)^2$;

$P_{RT}=(R+T)(S-U)/(2+R-T)^3$; $P_{SU}=0$.

$P_{RR}=(-2(-1+T)(S-U))/(2+R-T)^3$; $P_{SS}=0$

$P_{TT}=(-2(1+R)(S-U))/(2+R-T)^3$; $P_{UU}=0$

$Q_{RR}=(2(S-U))/(2+R-T)^3$; $Q_{SS}=0$

$Q_{TT}=(2(S-U))/(2+R-T)^3$; $Q_{UU}=0$

Table 7.1

Then

$P_r = aP_R + cP_S$, $\quad P_s = bP_R + dP_S$

$P_t = eP_T + gP_U$, $\quad P_u = fP_T + hP_U$.

$P_{rs} = abP_{RR} + (ad+bc)P_{RS} + cdP_{SS}$

$P_{rt} = aeP_{RT} + agP_{RU} + ceP_{ST} + cgP_{SU}$

$P_{ru} = afP_{RT} + ahP_{RU} + cfP_{ST} + ahP_{SU}$

$P_{st} = beP_{RT} + bgP_{Ru} + deP_{ST} + dgP_{SU}$

$P_{su} = bfP_{RT} + bhP_{RU} + dfP_{ST} + dhP_{SU}$

$P_{tu} = efP_{TT} + (eh+gf)P_{TU} + ghP_{UU}$

Set

$\chi = (2+R-T)$; $\quad \eta = S-U$; $\quad \zeta = 1+R$

$\omega = 1-T$.

The functions χ, η, and ζ are independent, and $\chi = \zeta + \omega$. It is easy to compute each of the expressions $\chi^2 P_r, \ldots, \chi^2 P_u, \chi^3 P_{rs}, \ldots, \chi^3 P_{tu}$, using Table 7.1. Denote by W(r,t,s,u)[i,j] the determinant formed by the i^{th} and j^{th} columns of the matrix W(r,t,s,u). It follows that

$\chi^5 W(r,t,s,u)[1,2]=$

$-(adg\chi^3) + a(de + bg)\chi^2\eta - abe\chi\eta^2 +$

$d(-ce + ag)\chi^2\zeta + b(ce - ag)\chi\eta\zeta$

$\chi^5 W(r,t,s,u)[1,3]=$

$agh\chi^3 + a(-(fg) - eh)\chi^2\eta + aef\chi\eta^2 +$

$h(ce - ag)\chi^2\zeta + f(-(ce) + ag)\chi\eta\zeta$

$\chi^6 W(r,t,s,u)[2,3]=$

$-(abgh\chi^4) + ab(fg + eh)\chi^3\eta - abef\chi^2\eta^2 +$

$(ag(bh-df)+bh(-ce + ag))\chi^3\zeta +$

$(bcef + adef - abfg - abeh)\chi^2\eta\zeta +$

$(ce-ag)(bh-df)\chi^2\zeta^2$

$\chi^5 W(s,u,r,t)[1,2]=$

$-(bch\chi^3) + b(cf + ah)\chi^2\eta - abf\chi\eta^2 +$

$c(-(df) + bh)\chi^2\zeta + a(df - bh)\chi\eta\zeta$

$\chi^5 W(s,u,r,t)[1,3]=$

$bgh\chi^3 -b(fg+eh)\chi^2\eta + bef\chi\eta^2 +$

$g(df - bh)\chi^2\zeta + e(-(df) + bh)\chi\eta\zeta.$

Table 7.2

Because the 2 x 2 subdeterminants of W(r,t,s,u) and W(s,u,r,t) must vanish identically on a neighborhood of the origin in $R^2 \times R^2$, each coefficient of a monomial in the variables χ[9], η, ζ that appears in Table 7.2 must be zero. Also note, that if α, β, ϵ, and γ are

9) This rescaling of the variables may involve rescaling the radius of convergence of the modules in the network.

nonzero real numbers, then

$$P(\ R,S,T,U\ )=C(\ A(\ r,t\ ),B(\ s,u\ )\ )$$

if and only if

$$P(\ \alpha R,\beta S,\epsilon T,\gamma U\ )=C''(\ A''(\ r,t\ ),B''(\ s,u\ )\ )$$

for some C",A", and B". This implies that one can, without loss of generality, multiply the rows of the change of coordinate matrices

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} e & f \\ g & h \end{vmatrix}$$

by nonzero constants when it is convenient to do so.

We are now in a position to analyze the derivatives in Table 7.2.

The coefficient of χ^3 in the expression $\chi^5 W(r,t,s,u)[1,2]$ is -adg. Similarly, each of the determinants that appears in Table 7.2 has a monomial in χ,η,ζ,ω that is a monomial in a,b,c,d,e,f,g, and h. Collect the monomial expressions into the following sets of equations. Each line in Table 7.3 corresponds to an equation in Table 7.2.

adg=0	
agh=0	aef=0
abgh=0	abe=0
bch=0	abf=0
bgh=0	bef=0.

Table 7.3

If $a=0$, because $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$

must have nonzero determinant, we can assume that b and c are not zero. Divide the first row by b and the second row by c. Therefore, suppose that $b=c=1$. But $bch=0$, therefore $h=0$. Thus, dividing the first row of

$$\begin{vmatrix} e & f \\ g & h \end{vmatrix}$$

by f and the second row by g, we can suppose that $f=g=1$. But $bef=0$, therefore $e=0$. The coefficient of the monomial $\chi^2\eta$ in $\chi^5 W(s,u,r,t)[1,2]$ is $b(cf + ah)$. But $ah=0$, therefore $bcf=0$. However, this contradicts the equation $f=1$. Therefore $a \neq 0$.

If we assume that $a \neq 0$, then we can divide the expression for R by a and assume that $a=1$. The equations in Table 7.3 imply that the following equations are satisfied

$dg=0$	
$gh=0$	$ef=0$
$bgh=0$	$be=0$
$bch=0$	$bf=0$
$bgh=0$	$bef=0$.

Table 7.4

If $b=0$, then because $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$

has nonzero determinant, we can assume that $a=d=1$. The first entry in Table 7.4 then shows that $g=0$, and therefore we can assume that e=h=1. It follows from the third line of Table 7.4 that ef=0, therefore f=0. The coefficient of $\chi^2\eta$ in $\chi^5 W(r,t,s,u)[1,2]$ is a(de + bg). Thus 0=ade=1.1.1, which is a contradiction. Therefore we can assume that a=1, b≠0. But the equations on the third and fourth lines of Table 7.4 are be=0 and bf=0. Then the matrix

$$\begin{vmatrix} e & f \\ g & h \end{vmatrix}$$

is singular. This contradicts the assumption that T=et+fu, U=gt+hu is a change of coordinates in the second agent's parameter space. We have shown that there are no linear coordinate changes A_1 and A_2 that can decrease the time required to compute P to two units of time.▒

There remains the possibility that simultaneous linear changes of coordinates in the agent's parameter spaces and in the message space could reduce the computing time. This is also not possible. The proof of this fact introduces nothing new and follows the pattern of the proof of Lemma 7.2.

Section III.

The Efficient Frontier

We examine two performance standards and analyze the efficient frontier for each. The two standards are each defined on the product of two two-dimensional Euclidean spaces. One performance standard is the function $I:R^2 x R^2 ---> R$, given by $I(u,v)=u \cdot v= I((x,z),(x',z'))=x x'+z z'$, the inner product. The second performance standard is the function $Q=(z-z')/(x-x')$. In the case of the function I, we show that there is a mechanism with message space of minimum dimension and with outcome function a projection that can be computed in minimum time and that realizes I. In the case of the function Q we show that if the dimension of the message space is allowed to increase, then the time required to compute the message correspondence μ, of Section I, can be reduced to two units of time. Recall that in this chapter the coordinate changes allowed are all linear. (In Chapter VIII we study the effect of analytic coordinate changes.)

We consider first the function I. It is well known (c.f. [11]) that the parameter transfer mechanism (with message space of dimension 3) has a message space of minimum dimension for mechanisms that realize I.

Suppose that $X=R^2$, $Y=R^2$, that X has coordinates (x,z), and that Y has coordinates (x′,z′). Suppose that the message space R^3 has coordinates (A,B,C). Then

$$I(x,z,x',z')= x\, x'+z\, z'.$$

A message correspondence for parameter transfer is given by the function

$\nu(x,z,x',z')=(x,z,I(x,z;x',z'))$. The agent with parameter space X uses as his message correspondence

$$\nu^1(x,z)=\{ (x,z,C): C\in R\},$$

while the agent with parameter space Y uses the correspondence

$$\nu^2(x',z')=\{ (A,B,C): A,B \in R,\ C=I(A,B, x',z')\}.$$

A network that computes the correspondence ν need only compute the function I(x,z,x′,z′) from the parameters x, z,x′,z′. This function I is a function of four variables that can be computed in two units of time by the network given in Figure 7.6. Thus, among mechanisms that realize I with outcome functions that are projections, no increase in the size of the message space will decrease the amount of computing required, since each such computation of a message correspondence must also compute I. It follows that the efficient frontier for the function I is given by the diagram in Figure 7.7.

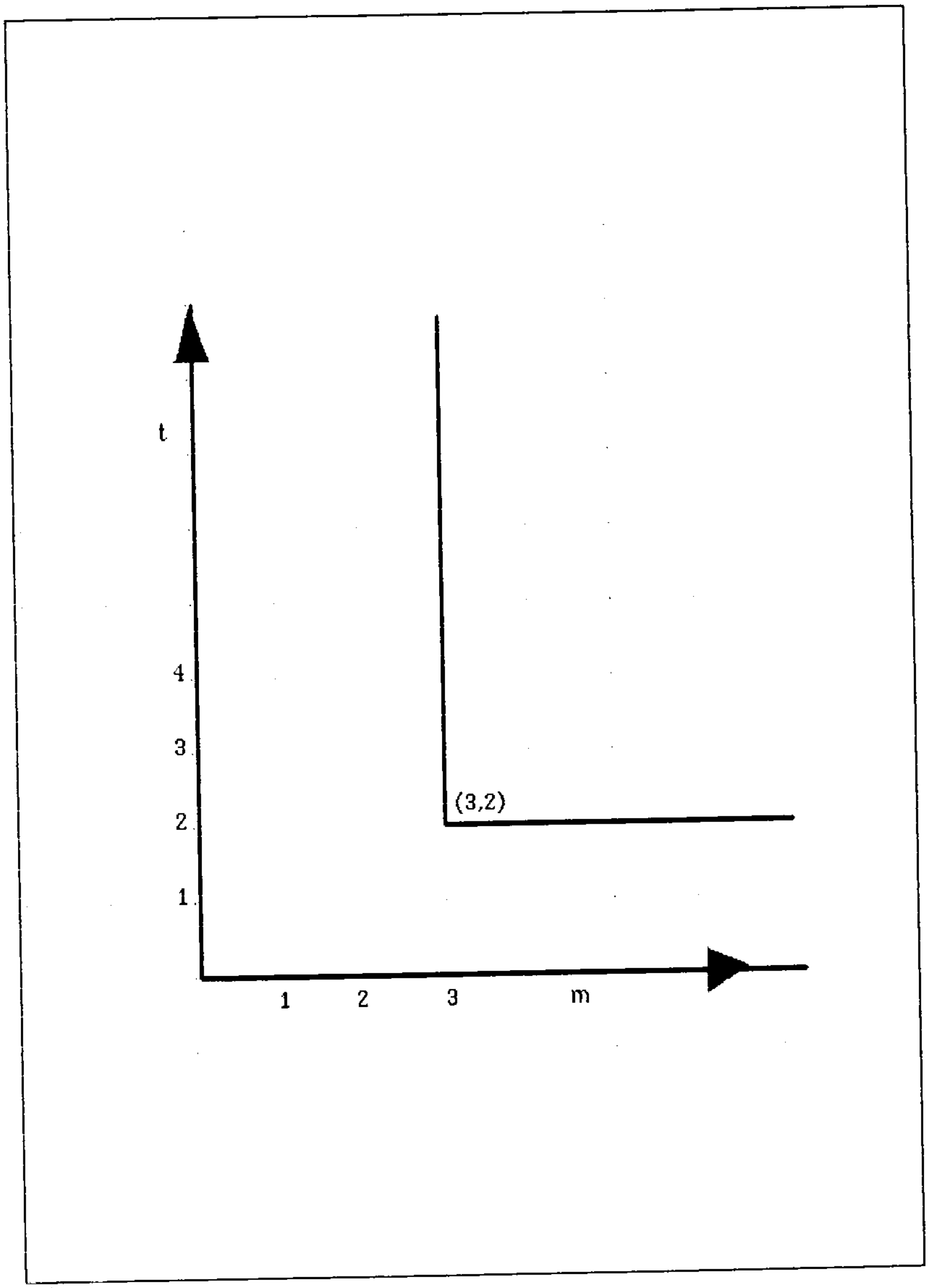

Figure 7.6

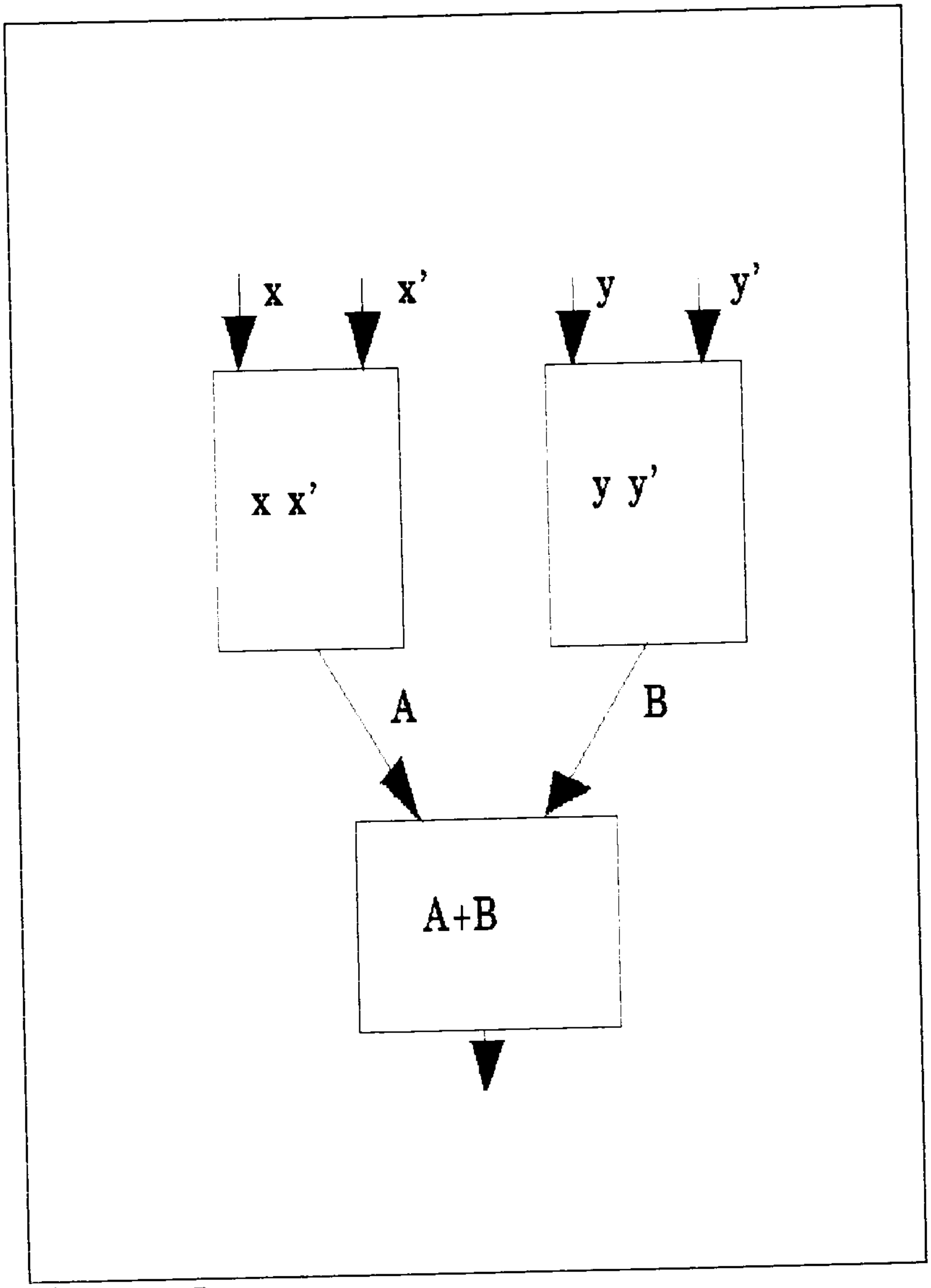

Figure 7.7

We turn now to the discussion of the function Q. Note that the function Q can be realized by the parameter transfer mechanism with R^3 as the message space. In that case Agent 1 has as message correspondence

$$\nu^1(x,z)=\{(X,Y,Z)|X=x, Y=z\}$$

while Agent 2 uses as message correspondence

$$\nu^2(x',z')=\{(X,Y,Z)|Z=(Y-z')/(X-x')\}.$$

The message correspondence for the mechanism is then

$$\nu(x,z;x',z')=(x,z,(z-z')/(x-x')).$$

Computing the function $(z-z')/(x-x')$, which is the only computation needed, requires two units of time using the network that is given in Figure 7.1.

Thus we see that increasing the dimension of the message space from 2 to 3 permits a decrease in the time required to compute the message correspondence from 3 units of time to 2 units of time. Because the minimum message space for Q is 2, the efficient frontier contains the points a and b shown in Figure 7.8.

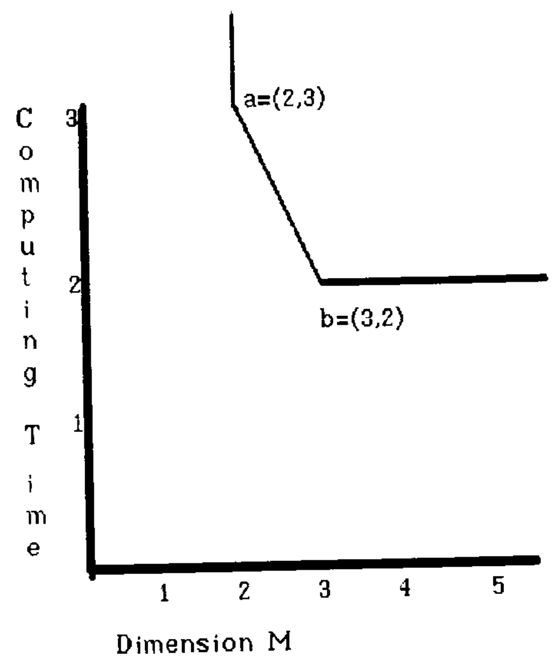

Figure 7.8

This means that the efficient frontier contains the points a=(2,3) and b=(3,2) shown in Figure 7.8. Because the minimum dimension for a message space of a privacy preserving mechanism that realizes Q is known to be 2, decreasing the dimension of the message space below 2 is impossible. Increasing the dimension of the message space above 3 does not yield any further decrease in the time required to compute the equilibrium message, because two units of time are required to compute Q, and Q is the projection of one of the coordinates of the message space. Therefore, the efficient frontier is as shown in Figure 7.8.

Remark. If we adopt the verification scenario as the interpretation of the mechanism, the computational task is:

(i) each agent computes his verifier function given a candidate equilibrium message,

(ii) the verification function is computed.

The computing time required for each of the agents to verify that a pair of values (Q,P) satisfies the equilibrium conditions

$P+xQ=z$ (for Agent 1)

and

$P+x'Q=z'$ (for Agent 2)

is the minimum possible for a function of 4 variables; that is, each requires two units of time. The computation of the function (Q,P) requires 3 units of time. Although the computational burden of computing the message μ(x,z;x',z') is decreased by increasing the dimension of the message space, this decrease is at the expense of an increased computational burden on the second agent. To check the equation $Z=(z-z')/(x-x')$ the agent must compute the function $Z-(z-z')/(x-x')$, and this is a function of 5 variables. The minimum computing time for such a function is $\text{Int}[\log_2(5)]=3$. This increase in computing time for agents is a feature of all the mechanisms we have examined.

Chapter VIII

The Effect of Automorphisms on Computational Complexity of an Edgeworth Box Economy with a Walrasian Performance Function

In Chapter VII we studied the efficient frontier for the performance standard

$$Q(x,z,x',z')=(z-z')/(x-x').$$

It was shown that for the mechanism

$$\mu(x,z,x',z')=\left(\frac{z-z'}{x-x'} , \frac{xz'-x'z}{x-x'} \right),$$

one cannot decrease the computing time required to compute the message $\mu(x,z,x',z')$ by applying linear nonhomogeneous coordinate transformations in the spaces of the agent's parameters or in the massage space. If we remove the restriction that the coordinate transformations must be linear, then the computing time for the message correspondence might be reduced by changing the coordinates in the agent's parameter spaces and changing the coordinates in the message space. There is also the possibility that there is a mechanism that realizes Q, that has a message correspondence that requires less than three units of time to compute, and cannot be derived from μ by applying coordinate changes in the message space and in the agent's parameter spaces. This second possibility, that there is a totally new mechanism, leads to the concept of isomorphism between mechanisms.

In Appendix A, there is a discussion of the concept of isomorphism of privacy preserving message correspondences when no topological conditions are placed on the functions and correspondences. However, the development can also be carried out when the performance standard is a continuous function and the message correspondence and outcome function are continuous. One can also consider isomorphisms when the maps and correspondences are differentiable. Roughly, two mechanisms are isomorphic if one mechanism can be transformed into the other by applying coordinate changes in each of the agent's parameter spaces and in the coordinate system chosen for the message space. In the first section of this chapter we show that, under some smoothness conditions, each mechanism that realizes the function

$$Q(x,z,x',z')=(z-z')/(x-x')$$

and that has a two dimensional message space is, locally, isomorphic to the mechanism given in Chapter VII. This result is closely related to the theorem of Jordan[13].

In the second section of this chapter we show that the time required to compute the message correspondence

$$\mu(x,z,x',z')=\left(\frac{z-z'}{x-x'} , \frac{xz'-x'z}{x-x'} \right),$$

cannot be reduced below three units of time by an analytic coordinate change in the message space when the outcome function is assume to be a projection. Finally, we show that real analytic coordinate transformations in the agents' parameter spaces cannot reduce the computation time required to compute the message correspondence for μ below three units. We have seen that Q can be computed in two units of time with a proper choice of coordinates. Similarly, we will show that with a proper choice of coordinates P can be computed in two units of time . However, different coordinates are required in each case. Thus three units of computing time are required to compute both when a single coordinate system is used.

Section I

Local Isomorphism of Mechanisms Realizing Q

In this section we prove that, to within local isomorphism, there is only one mechanism that realizes Q using a two dimensional message space and a message correspondence that is a function. The concept of isomorphism is the topological version (c.f. Appendix A.)

Lemma 8.1. Assume that V_1 and V_2 are nonempty open subsets of R^2 and assume that $Q:R^2 \times R^2 \text{--->} R$ is the function

$$Q(x,z,x',z')=(z-z')/(x-x').$$

Suppose that

$$\Delta=\{(x,z,x',z') \mid (x-x')=0\}.$$

Assume that $m:V_1xV_2-\Delta$ --->M is a privacy preserving correspondence to a Euclidean space M that satisfies the following conditions:

(i) M is a two dimensional Euclidean space,

(ii) there is a submersion h:M--->R such that the triple (ν,M,h) realizes Q,

(iii) the function ν is a differentiable function that is a submersion on $V_1xV_2-\Delta$,

(iv) the coordinate correspondences $\nu_i:V_i$--->M, i=1,2, are nonsingular correspondences, i.e. the correspondences (as subsets of V_ix M) are submanifolds, the projection of ν_i onto M is a submersion and the sets $\nu_i^{-1}(p)$ are nonsingular submanifolds of $V_1xV_2-\Delta$,

(v) for each $p \in M$, and each i, the set $\nu_i^{-1}(p)$ is a nonsingular submanifold of $V_1 x V_2 - \Delta$ of dimension d_i (independent of p).

Suppose $m_0 \in M$, and suppose $a_0 \in V_1$, and $a'_0 \in V_2$. In a neighborhood of the point m_0, there is a coordinate system (S,T) and a choice of coordinates (ξ, ζ) in a neighborhood of a_0 and a choice of coordinates (χ', ζ') in a neighborhood of a'_0 so that the correspondence ν is the function

$$\nu(\chi, \zeta, \chi', \zeta') = \left(\frac{\zeta - \zeta'}{\chi - \chi'} , \frac{\chi \zeta' - \chi' \zeta}{\chi - \chi'} \right)$$

Proof. Suppose that M has coordinates (M_1, M_2) and suppose that $(x_0, z_0) = a_0$, $(x'_0, z'_0) = a'_0$. Set

$$c = h \cdot m(x_0, z_0, x'_0, z'_0).$$

Therefore

$$Q(x_0, z_0, x'_0, z'_0) = c$$

and

$$Q^{-1}(c) =$$

$$\{(x,z,x',z') | (z-z') - c(x-x') = 0\} \cap (V_1 \text{ x } V_2 - \Delta).$$

Set

$$\nu(x_0, z_0, x'_0, z'_0) = (p_1, p_2).$$

It follows from condition (ii) that the function $h(M_1, M_2) - c$ is a nonsingular function on M that is zero at (p_1, p_2). We can find a function h':M--->R that is differentiable on a neighborhood of (p_1, p_2) so that the pair (h-c,h') is a local coordinate system on M.

The function $H^*=(h-c,h')$ carries a neighborhood of (p_1,p_2) to an open neighborhood of the origin of R_2. The function $(h-c)\cdot\nu=Q$. Set $f=h'\cdot\nu$. Replace M by a neighborhood U of the origin of R_2 and replace the function ν by the function $\nu^*=(Q-c,f)$. Denote the coordinates functions on U by X and Y. Set $\nu^*_i=H^*\cdot\nu_i$. Because ν is a privacy preserving correspondence that realizes Q, it follows that $\nu^{*-1}(0,0)$ is a rectangle in $\nu^{-1}(c)$ that contains the point $(x_0,z_0; x'_0,z'_0)$. Assumption (ii) implies that the set $\nu^{*-1}(0,0)$ is a nonsingular submanifold of the set $V_1 \times V_2-\Delta$. Because ν is privacy preserving, it follows that there are correspondences ν^*_1 and ν^*_2 such that $\nu^*=\nu^*_1\cap\nu^*_2$. Then $\nu^{*-1}(0,0)$ is the product $C_1 \times C_2$ where $C_1=\nu^*_1{}^{-1}(0,0)$ and $C_2=\nu^*_2{}^{-1}(0,0)$. Each $\nu_i{}^{-1}(0,0)$ is a nonsingular submanifold of $V_1 \times V_2-\Delta$, by assumption (iv). Furthermore, $\nu^{*-1}(0,0)=Q^{-1}(c)$, and the set $Q^{-1}(c)$ is a submanifold of $V_1 \times V_2-\Delta$ of dimension 3. The restriction of ν^* to the set $\nu^{*-1}(U)\cap Q^{-1}(c)$ carries $\nu^{*-1}(U)\cap Q^{-1}(c)$ onto the set $U\cap\{(X,Y)\,|\,X=0\}$. The mapping H^* is a homeomorphism on a neighborhood of p, therefore the restriction to U of the correspondences ν^*_i also have that are nonsingular submanifolds of $R^2 \times M$. Because ν^* is a submersion, for each point $p\in M$, the dimension of the submanifold $\nu^{*-1}(p)$ is 2. Therefore, the dimension of the submanifold $\nu*_i^{-1}(p)$ is either 0, 1, or 2.

Suppose that the dimension of $\nu*_1^{-1}(0,0)$ is 2. The point $(x_0,z_0;x'_0,z'_0)$ is in $\nu*^{-1}(p)$. Each point $(x,z;x'_0,z'_0)$ must also be in the rectangle $\nu*^{-1}(p)$, for each $(x,z)\in R^2$ such that $x\neq x'_0$. But this implies that

$$Q(x,z;x'*,z'*)=(z-z'_0)/(x-x'_0)$$

is independent of x and z. Since this is clearly impossible, we can assume that the dimension of $\nu*_i^{-1}(p)$ is 1 for each $p\in U$.

We denote by $\nu*_1$ the restriction of ν_1 to the set V_1 x U (that is $\nu*_1=\nu_1\cap(V_1 x U)$). Then the projection from $\nu*_1$ to U is a submersion. Furthermore, for each $p\in U$, the set $\nu*_1^{-1}(p)$ is a submanifold of V_1 of dimension 1.

If the dimension of $\nu*_1$ is 2, because the projection pr_1 to U is a submersion, the map pr_1 would be a bijection [c.f. 7, p. 7], and $\nu*_1^{-1}(p)$ would be zero dimensional. This is impossible. Therefore $\nu*_1^{-1}(p)$ has dimension greater than 2.

Because the values of Q depend on the parameters of the first agent, the dimension of $\nu*_1$ cannot be 4. Therefore, the dimension of $\nu*_1$ is 3.

We may suppose that $\nu*_1$ has (for a sufficiently small open set that contains $(x_0,z_0;0,0)$) an equation

$F(x,z;X,Y)=0$, where G is a C_2 function. The set in V_1 with equation $F(x,z;0,0)=0$ is the submanifold $\nu_1^{-1}(0,0)$ that has dimension 1. Because $\nu_1^{-1}(0,0)$ is a nonsingular curve in V_1, it follows that one of the partial derivatives $\partial F/\partial x$ or $\partial F/\partial z$ is nonzero at $(x_0,z_0;0,0)$. Furthermore, because $\partial Q/\partial z\neq 0$, we can assume that $\partial F/\partial z\neq 0$. In a neighborhood of $(x_0,z_0;0,0)$ the solution of the equation $F(x,z;X,Y)=0$ is a function $f(x;X,Y)$ such that $F(x,f(x,X,Y);X,Y)=0$. For each X and Y in a sufficiently small neighborhood of (0,0), the function $(x,f(x;X,Y))$ parameterizes the curve $\nu_1^{-1}(X,Y)$. Similarly, we assume that $G(x',z';X,Y)=0$ is an equation for the correspondence ν_2 in a neighborhood of $(x'_0,z'_0;0,0)$. One of the derivatives $\partial G/\partial x'$ or $\partial G/\partial z'$ is nonzero at $(x'_0,z'_0;0,0)$. Assume that $\partial G/\partial z'$ is not zero. Solve the equation $G(x',z';X,Y)=0$, for z' in a sufficiently small neighborhood of the origin. That is, there is a function $g(x';X,Y)$ so that $(x',g(x';X,Y))$ parameterizes the curve $m*_2^{-1}(X,Y)$. It follows that for each X sufficiently close to 0, the points $(x,f(x;X,Y),x',g(x',X,Y))$ are in the set with equation $Q(x,z;x',z')-X=0$. Therefore

$$(f(x;X,Y)-g(x';X,Y))/(x-x')=X.$$

That is,

$$f(x;X,Y)-g(x';X,Y)=X(x-x'),$$

or

$f(x;X,Y)-xX=g(x';X,Y)-x'X$. Because the right hand side of this equation is independent of x', $f(x;X,Y)-xX$ is a function independent of x and x'. That is

$$f(x;X,Y)-xX=K(X,Y).$$

The function $F(x,z;X,Y)$ has partial derivative $\partial F/\partial Y\neq 0$ in a neighborhood of $(0,0)$, because the restriction of pr_M, the projection of $V_1 \times V_2 \times M$ to ν_1, was assumed to be a submersion. But

$$\partial F/\partial Y+(\partial F/\partial z)(\partial f/\partial Y)=0.$$

Therefore $\partial f/\partial Y\neq 0$ at $(0,0)$ for x sufficiently close to 0. Because

$$K(X,Y)=f(x,X,Y)-xX,$$

it follows that $\partial K/\partial Y\neq 0$ at $(0,0)$. We introduce as new coordinates on M, in a neighborhood of the point $(0,0)$, the pair of functions X and $K(X,Y)$. The function Y satisfies a relation $Y=K^*(X,K)$ in a neighborhood of $(0,0)$. Therefore

$$f(x,X,Y)=f(x,X,K^*(X,K)).$$

For each X and K, the pair $(x,f(x,X,K^*(X,K)))$ parameterizes the curve $\nu^*{}_1{}^{-1}(X,K^*(X,K))$. We now wish to find the expression for the function ν^* in the new coordinates for M. For a point $(x,z;x',z')$ in $V_1 \times V_2$ such that $Q(x,z;x',z')=X$, suppose that $\nu(x,z;x',z')=(X,k)$. Thus $(x,z;x',z')$ lies on the product of the curves $\nu^*_1{}^{-1}(x,z)$ and $\nu^*_2{}^{-1}(x',z')$. Therefore,

$$f(x,X,K^*(X,k))=z$$

and

$$g(x',X,K^*(X,k))=z'.$$

The point (X,k) must be on the intersection of the two curves in M with equations $z-xX=k$ and $z'-x'X=k$. That is, the value for k satisfies the equations $k+xX=z$ and $k+x'X=z'$. If we solve these equations,

$$X=(z-z')/(x-x')$$

and

$$k=(zx'-xz')/(x-x'). ▒$$

Section II

The Effect of Automorphisms on Computing Time

In this section we show that no automorphism of the message space RxR can be composed with

$$\mu=(Q,P)=((z-z')/(x-x'),(zx'-xz')/(x-x')),$$

and produce a message correspondence that can be computed in less than 3 units of time.

<u>Theorem 8.1</u> Suppose

$$Q=(z-z')/(x-x')$$

and

$$P=(xz'-x'z)/(x-x').$$

If $F=A(Q,P)$ and $Q=B(Q,P)$ and if the map (u,v)--->(A(u,v),B(u,v)) is a C^2 automorphism of RxR, then each circuit that computes both Q and F requires at least 3 units of computing time.

Proof. We have already noted in the proof of Theorem 7.1 that coordinate changes that are translations do not effect computation time. Therefore, without loss of generality, we introduce the following changes of coordinates in the agents parameters;

$x=R+1$, $z=S$, $z'=U$, and $x'=T-1$.

In the coordinates R, S, T, and U ,

$Q=(S-U)/(2+R-T)$

and

$P=(U+S+RU-ST)/(2+R-T)$.

Because the map M(Q,P)=(Q,F)=(Q,A(Q,P)) is an automorphism, $A_P \neq 0$ in a nonempty open set. The message function (Q,P) composed with M expresses Q and F as functions of R, S, T, and U. It follows from the Chain Rule that for X=R, S, T, or U,

$F_X=A_Q Q_X+A_P P_X$.

Therefore we have the following expression for the second derivative in X and Y, when X and Y are chosen from the set R, S, T, U:

$F_{XY}=A_{QQ}Q_X Q_Y+A_{PQ}P_Y Q_X+A_Q Q_{XY}+A_{PQ}Q_Y P_X+A_{PP}P_Y P_X+A_P P_{XY}$.

Table 7.1 of Chapter VII lists the derivatives of the functions Q and P. In order that the function F be computable in time 2, the matrix

$$\text{(I)} \quad \begin{array}{c|ccc|} & 1 & R & T \\ S & F_S & F_{RS} & F_{ST} \\ U & F_U & F_{RU} & F_{TU} \end{array}$$

and the matrix

$$\text{(II)} \quad \begin{array}{c|ccc|} & 1 & S & U \\ R & F_R & F_{RS} & F_{RU} \\ T & F_T & F_{ST} & F_{TU} \end{array}$$

must have rank at most 1, or the matrices

$$\text{(III)} \quad \begin{array}{c|ccc|} & 1 & R & U \\ S & F_S & F_{RS} & F_{SU} \\ T & F_T & F_{RT} & F_{TU} \end{array}$$

and

$$\text{(IV)} \quad \begin{array}{c|ccc|} & 1 & S & T \\ R & F_R & F_{RS} & F_{RT} \\ U & F_U & F_{SU} & F_{TU} \end{array}$$

both must have rank at most 1. As in Chapter VII, set

$\chi=(2+R-T)$; $\eta=S-U$; $\zeta=1+R$;

$\omega=1-T$. Note, as we did before, that the functions χ, η, and ζ are independent and that $\chi=\zeta+\omega$. Table 8.1 presents a matrix M with rows indexed by products XxY, where X and Y are chosen from the set {R,S,T,U}. The columns of M are products of P and Q. The entry in the row (X x Y) and column (AxB) is the product $A_X B_Y$

expressed in terms of the parameters χ, η, ζ, and ω. For example the entry in row (SxR) and column PxQ is

$$P_S\ Q_R=[(1-T)/(2+R-T)][-(S-U)/(2+R-T)^2]$$

$$=[\omega/\chi][-\eta/\chi^2]=-\omega\eta/\chi^3.$$

	P	Q	PxP	PxQ	QxP	QxQ
R	$-\omega\eta/\chi^2$	$-\eta/\chi^2$	0	0	0	0
S	ω/χ	$1/\chi$	0	0	0	0
T	$-\zeta\eta/\chi^2$	η/χ^2	0	0	0	0
U	ζ/χ	$-1/\chi$	0	0	0	0
RxR	$2\omega\eta/\chi^3$	$2\eta/\chi^3$	$\omega^2\eta^2/\chi^4$	$\omega\eta^2/\chi^4$	$\omega\eta^2/\chi^4$	η^2/χ^4
RxS	$-\omega/\chi^2$	$-1/\chi^2$	$-\omega^2\eta/\chi^3$	$-\omega\eta/\chi^3$	$-\omega\eta/\chi^3$	$-\eta/\chi^3$
RxT	$(\zeta-\omega)\eta/\chi^3$	$-2\eta/\chi^3$	$\omega\zeta\eta^2/\chi^4$	$-\omega\eta^2/\chi^4$	$\zeta\eta^2/\chi^4$	$-\eta^2/\chi^4$
RxU	$-\omega/\chi^2$	$1/\chi^2$	$-\omega\zeta\eta/\chi^3$	$\omega\eta/\chi^3$	$-\zeta\eta/\chi^3$	η/χ^3
SxR	$-\omega/\chi^2$	$-1/\chi^2$	$-\eta\omega^2/\chi^3$	$-\omega\eta/\chi^3$	$-\omega\eta/\chi^3$	$-\eta/\chi^3$
SxS	0	0	ω/χ^2	ω/χ^2	ω/χ^2	$1/\chi^2$
SxT	$-\zeta/\chi^2$	$1/\chi^2$	$-\zeta\eta\omega/\chi^3$	$\omega\eta/\chi^3$	$-\zeta\eta/\chi^3$	η/χ^3
SxU	0	0	$\zeta\omega/\chi^2$	$-\omega/\chi^2$	ζ/χ^2	$-1/\chi^2$
TxR	$(\zeta-\omega)\eta/\chi^3$	$-2\eta/\chi^3$	$\zeta\omega\eta^2/\chi^4$	$\zeta\eta^2/\chi^4$	$-\omega\eta^2/\chi^4$	$-\eta^2/\chi^4$
TxS	$-\zeta/\chi^2$	$1/\chi^2$	$-\zeta\eta\omega/\chi^3$	$-\zeta\eta/\chi^3$	$\eta\omega/\chi^3$	η/χ^3
TxT	$-2\zeta\eta/\chi^3$	$2\eta/\chi^3$	$\zeta^2\eta^2/\chi^4$	$-\zeta\eta^2/\chi^4$	$-\zeta\eta/\chi^4$	η^2/χ^4
TxU	ζ/χ^2	$-1/\chi^2$	$-\zeta\eta^2/\chi^3$	$\zeta\eta/\chi^3$	$\zeta\eta/\chi^3$	$-\eta/\chi^3$
UxR	$-\omega/\chi^2$	$1/\chi^2$	$-\zeta\omega\eta/\chi^3$	$-\zeta\eta/\chi^3$	$\omega\eta/\chi^3$	η/χ^3
UxS	0	0	$\zeta\omega/\chi^2$	ζ/χ^2	$-\omega/\chi^2$	$-1/\chi^2$
UxT	ζ/χ^2	$-1/\chi^2$	$-\zeta^2\eta/\chi^3$	$\eta\zeta/\chi^3$	$\eta\zeta/\chi^3$	$-\eta/\chi^3$
UxU	0	0	ζ^2/χ^2	$-\zeta/\chi^2$	$-\zeta/\chi^2$	$1/\chi^2$

Table 8.1

The Chain Rule shows that the vector $(F_R, F_S, F_T, F_U, F_{RR}, F_{RS}, \ldots, F_{SR}, \ldots, F_{UU})^T$, where the superscript T denotes the transpose, is the product $M\ (A_P, A_Q, A_{PP}, A_{PQ}, A_{QP}, A_{QQ})^T$. Now set $\eta=0$ and evaluate the matrices (I) and (II). The matrix (I)=

$(\omega/\chi)A_P+(1/\chi)A_Q \quad (-\omega/\chi^2)A_P-(1/\chi^2)A_Q \quad (-\zeta/\chi T2)A_P$

$(\zeta/\chi)A_P-(1/\chi)A_Q \ (-\omega/\chi^2)A_P-(1/\chi^2)A_Q \ (\zeta zgw/\chi^2)A_{PP}+(\zeta-\omega)/\chi^2 A_{PQ}-1/\chi^2 A_{QQ}$

and the matrix (II)=

$$\begin{matrix} 0 & (-\omega/\chi^2)A_P-(1/\chi^2)A_Q & (-\omega/\chi^2)A_P+(1/\chi^2)A_Q \\ 0 & (-\zeta/\chi^2)A_P+(1/\chi^2)A_Q & (\zeta/\chi^2)A_P-(1/\chi^2)A_Q. \end{matrix}$$

If (I) has rank less than 2, then

$(\omega A_P+A_Q)(\omega+\zeta)A_P=0$

and if (II) has rank at most 1, then

$(\zeta A_P-A_Q)(-2\omega A_P)=0.$

In the set where $\omega\neq 0$, because $A_P\neq 0$,

$A_Q=\zeta A_P.$

But then,

$(\omega+\zeta)^2A_P=0.$ However, this is impossible.

Thus if F can be computed in two units of time, the matrices (III) and (IV) must have rank at most 1. It is easy to see that if (III) and (IV) have rank at most one, then either $A_Q=-\omega A_P$, or $\zeta A_P=A_Q$. But then $A_Q=0$ when $\omega=0$ or $A_Q=0$ if $\zeta=0$. But $A_Q\neq 0$ in a nonempty open set, therefore (III) and (IV) cannot both have rank at most 1. It follows that F cannot be computed in 2 units of time. ▒

We have shown that no automorphism of the message space can decrease the computing time below

three units of time when the outcome function is a projection. There is still the possibility that changes in the coordinates of the agent's parameter spaces can decrease the time required for the computation of the message correspondence μ. Lemma 8.2 addresses that possibility.

<u>Lemma 8.2.</u> If $m:R^2 \times R^2 \text{--->} R^2$ is the function given by $m(x,z;x',z')=(Q,P)$, if

$Q=(z-z')/(x-x')$

and

$P=(xz'-x'z)/(x-x')$,

and if coordinate systems in R^2 are chosen so that one of the two functions Q or P can be computed in two units of time using functions of two variables that are nonsingular, then in those coordinates the other function requires at least three units of computation time.

Proof. Suppose that we are using the representation of the functions Q and P used in Theorem 8.1. In the coordinates R, S, T, U,

$Q=(S-U)/(2+R-T)$

and

$P=(U+S+RU-ST)/(2+R-T)$.

Suppose that Q can be computed in two units of time, using coordinates (r,s) in the (R,S) space and coordinates (t,u) in the (T,U) space.

Notation. Set $X_{ab} = \partial^2 X / \partial a \partial b$ for a function X of variables a and b.

If

Q=C(A(r,t),B(s,u))

and

P=C'(A'(r,t), B'(s,u)),

then the criteria given in Theorem 6.1 of Chapter VI shows that each of the following matrices must have rank at most 1.

$$\text{M1:}\begin{vmatrix} Q_r & Q_{rs} & Q_{ru} & Q_{rss} & Q_{rsu} & Q_{ruu} \\ Q_t & Q_{st} & Q_{tu} & Q_{tss} & Q_{tsu} & Q_{tuu} \end{vmatrix}$$

$$\text{N1:}\begin{vmatrix} Q_r & P_{rs} & P_{ru} & P_{rss} & P_{rsu} & P_{ruu} \\ P_t & P_{st} & P_{tu} & P_{tss} & P_{tsu} & P_{tuu} \end{vmatrix}$$

$$\text{MI:}\begin{vmatrix} Q_s & Q_{rs} & Q_{st} & Q_{rrs} & Q_{rst} & Q_{stt} \\ Q_u & Q_{ru} & Q_{tu} & Q_{rru} & Q_{rtu} & Q_{ttu} \end{vmatrix}$$

$$\text{NI:}\begin{vmatrix} P_s & P_{rs} & P_{st} & P_{rrs} & P_{rst} & P_{stt} \\ P_u & P_{ru} & P_{tu} & P_{rru} & P_{rtu} & P_{ttu} \end{vmatrix}$$

Furthermore, the matrices $\begin{vmatrix} R_r & R_s \\ S_r & S_s \end{vmatrix}$ and $\begin{vmatrix} T_t & T_u \\ U_t & U_u \end{vmatrix}$

must have nonzero determinants D and E, respectively. Because R and S are assumed to be functions of r and s alone, while T and U are assumed to be functions of t and u alone, it is tedious but elementary to show that the matrices Q1, P1, QI, and PI each has rank at most one only if there are real numbers K, L, M, and N so that the following conditions are satisfied.

Q1 : $(1/2)\ S_r = L\{(-1/2)\ U_t\}$

Q2 : $3\ S_{rs} -(1/2)[R_r\ S_s + S_r R_s] = L\{U_t R_s\}$

Q3 : $R_r U_u = L\{-3U_{tu}\}$

Q4 : $(3/2)S_{rss} -(1/2)[S_r R_{ss} + 2S_s\ R_{rs}] -(1/2)[R_r\ S_{ss} + 2\ R_s\ S_{rs}] = L\{(1/2)\ R_{ss}\ U_t\}$

Q5 : $R_{rs} U_u = L\{U_{tu}\ R_s\}$

Q6 : $(1/2)\ U_{uu}\ R_r = L\{(-3/2)\ U_{tuu}\}$

QI : $(1/2)S_s = M\{(-1/2)\ U_u\}$

QII : $3S_{rs} -(1/2)[R_r S_s + S_r R_s] = M\{R_r\ U_u\}$

QIII : $R_s\ U_t = M\{(-3)U_{tu}\}$

QIV : $(3/2)\ S_{rss} -(1/2)[S_s\ R_{rr} + 2\ S_r R_{rs}] = M\{(1/2)\ R_{rr}\ U_u\}$

QV : $R_{rs}\ U_t = M\{U_{tu}\ R_r\}$

QVI : $U_{tt}\ R_s = M\{(-3/2)\ U_{ttu}\}$

P1 : $1/2\ S_r = N\{(-1/2)U_t\}$

P2 : $3S_{rs}-(1/2)(S_rR_s+R_rS_s) = N\{-S_sU_t\}$

P3 : $3U_uR_r\ -T_uS_r = N\{(-3)U_{ut}-(1/2)[U_tT_u + T_tU_u]\}$

P4 : $(3/2)S_{rss}-(1/2)[R_rS_{ss}+2R_sS_{rs}]-$

$(1/2)[S_rR_{ss}+2S_sR_{rs}] =$

$N[-(1/2)S_{ss}T_t + (3/2)R_{ss}U_t]$

P5 : $-S_{rs}T_u + 3R_{rs}U_u = N\{3U_{tu}R_s - T_{tu}S_s\}$

P6 : $(3/2)U_{uu}R_r-(1/2)T_{uu}S_r = N\{(-1/2)[T_tU_{uu} +$

$2\ T_u\ U_{tu}] + (-1/2)[U_tT_{uu} + 2U_u\ T_{tu}]\}$

PI : $(1/2)S_s = K[(-1/2)\ U_u]$

PII : $3\ S_{rs}\ -(1/2)[S_rR_s + S_sR_r] = K\{3\ R_r\ U_u - T_u\ S_r\}$

PIII : $-T_tS_s = K\{(-3/2)\ U_{tu}\ -(1/2)[\ U_t\ T_u + T_t\ U_u]\}$

PIV : $(3/2)S_{rrs} - (1/2)[R_sS_{rr} + 2\ R_rS_{rs}] -$

$(1/2)[S_s\ R_{rr} +2\ S_r\ R_{rs}] = K\{(-1/2)\ S_{rr}\ T_u +$

$(3/2)\ R_r\ U_u\}$

PV : $-S_{rs}\ T_t + 3\ R_{rs}\ U_t = K\{3\ U_{tu}\ R_r- T_{tu}\ S_r\}$

PVI : $(3/2)\ U_{tt}\ R_s - (1/2)\ T_{tt}\ S_s = K\{(-3/2)U_{ttu} -$

$(1/2)[T_uU_{tt} + 2T_t\ U_{tu}] -(1/2)[U_u\ T_{tt} + 2\ U_t\ T_{tu}]\}$

Note that the sixteen equations X1, X2, X3, X5, XI, XII, XIII, XV, with X equal to Q or P, involve only the sixteen variables K, L, M, N, R_r, S_r, R_s, S_s, R_{rs}, S_{rs}, T_t, U_t, T_u, U_u, T_{tu}, and U_{tu}.

Suppose that L=0. Equation Q1 implies that

$S_r=0.$

But then P1 implies that

$NU_t=0.$

If both L=0 and N=0, then P3 implies that

$3U_u\ R_r = T_u\ S_r.$

Because $S_r=0$ and the determinant $D\neq0$, it follows that

$R_r\neq0,$

from which it follows that

$U_u=0.$

However, QI then implies that

$S_s=0.$

But then D=0, which contradicts the assumption $D\neq0$.

Therefore we can assume that

$L=0$, but $N\neq0$.

Then

$U_t=0.$

Equation Q2 implies that

$3S_{rs}=(1/2)\ R_rS_s$

and QII implies that

$3S_{rs}=(1/2)R_rS_s + M\ R_r\ U_u.$

Therefore

$$M\ R_r\ U_u = 0.$$

But if $S_r=0$ and $D\neq0$, then

$R_r\neq0$, while if $U_t=0$, then $U_u\neq0$.

It follows that M=0. However, QI then implies that $S_s=0$ which contradicts the assumption that D=0.

We conclude from this that L must be nonzero.

Suppose that $U_t=0$.

Then Q1 shows that

$$S_r=0.$$

Then Q2 shows that

$$3S_{rs} = (1/2)\ R_r\ S_s$$

while QII shows that

$$3S_{rs}=(1/2)R_r\ S_s+M\{R_r\ U_u\}.$$

Therefore,

$$M\ R_r\ U_u = 0.$$

But $R_r\neq0$ because

$$S_r=0,$$

and

$U_u\neq0$ because $U_t=0$.

Therefore M=0.

But then the equation QI shows that

$$S_s=0$$

which contradicts the assumption that $D\neq0$.

Thus we can assume that

$U_t \neq 0$ and $L \neq 0$.

Now suppose that

$M=0$.

It follows from QI that

$S_s=0$.

If we substitute 0 for S_s in Equation PI, it follows that

$K\ U_u=0$.

If we also assume that $K=0$,

then PII shows that

$3S_{rs}=(1/2)\ S_r\ R_s$

while Equation Q2 shows that

$3\ S_{rs}=(1/2)R_s\ S_r\ +L\ U_t\ R_s$.

Therefore

$L\ U_t\ R_s=0$.

But $L \neq 0$ and $S_s=0$ implies that

$R_s\ \neq 0$.

Therefore $U_t=0$.

But then Q1 shows that

$S_r=0$

which contradicts the assumption that $D \neq 0$.

Therefore if we assume that $M=0$ we must conclude that $K \neq 0$. However, it then follows that

$U_u=0$.

From QIII it follows that

$R_s\ U_t = 0.$

But $U_t \neq 0$ because $E \neq 0$ and

$U_u = 0$

while

$R_s \neq 0$

because

$D \neq 0$ and $S_s = 0$ by QI.

Thus we can conclude that $M \neq 0$.

Suppose that

$U_u = 0.$

Equation QI implies that

$S_s = 0.$

It follows from QII that

$3\ S_{rs} = (1/2)\ S_r\ R_s$

while Q2 implies that

$3S_{rs} = (1/2)\ S_r\ R_s + L(U_t\ R_s).$

Therefore

$L\ U_t\ R_s = 0.$

But $L \neq 0$ from which it follows that either

$R_s = 0$ or $U_t = 0.$

If $R_s = 0$ then $D = 0$

and if

$U_t = 0$ then $E = 0.$

Thus we can assume that $U_u \neq 0$.

When $LMU_t\ U_u \neq 0,$

then Q1 and P1 imply that

$L=N$,

while QI and PI imply that

$M=K$.

Furthermore, Q1 shows that

$S_r=-LU_t$

and QI shows that

$S_s=-MU_u$.

The equations Q2 and P2 show that

$L\ R_s\ U_t=LM\ U_u\ U_t$

from which we can conclude that

$R_s=M\ U_u$.

Equation Q3 can be solved to yield

$U_{tu}=(-1/3L)\ R_r\ U_u$

while QIII shows that

$U_{tu}=(-1/3)\ U_t\ U_u$.

It follows that

$L\ U_t\ U_u=R_r\ U_u$ and therefore $R_r=L\ U_t$.

Next we turn to equation Q5. If one substitutes the value computed for U_{tu} into this expression we can conclude that $R_{rs}=(-1/3)\ LM\ U_t\ U_u$.

We also find that value of the expression

$R_r\ S_s+S_r\ R_s=-2LM\ U_t\ U_u$.

From QII it follows that

$3\ S_{rs}=MLU_tU_u-LMU_t\ U_u=0$.

The equation QV shows that

$(-1/3)U_t^2\ U_u=(-1/3)ML\ U_t^2\ U_u$

and therefore that LM=1.

Substitution of the values we have computed for S_{rs} , R_{rs}, and S_s into equation P5 shows that

$$T_{tu}=0.$$

From equation PV we can then conclude that

$$T_t=U_t.$$

We can then solve equation P3 for T_u to find that

$$T_u=(-5/3)\ U_u.$$

But then PII shows that

$$LMU_t\ U_u=3\ LM\ U_t\ U_u-(5/3)U_t\ U_u,$$

which is clearly impossible.

We conclude that no analytic coordinate changes in the agents parameters can decrease the time required to compute a realization of Q using a message space of dimension 2, when the computation of Q as the outcome function is required. In the last paragraph of Chapter VII we also noted that in the case of the verification scenario, a central agent can compute in parallel both the messages that are to be sent to the agents for verification and the performance standard. It is conceivable that there is a mechanism realizing Q with a message space of dimension two that requires only two units of computation time because the computation of the outcome function is not required.

Computational Complexity of Mechanisms

Chapter IX

Separator Sets for Smooth Functions - I

We investigate the relations between computations using finite networks processing finite alphabets and computations using continuous networks as defined in Chapter III. We seek to clarify the sense in which the computation of a continuous function by a continuous network can be considered as a limit of computations by finite networks computing finite approximations to the continuous function. Chapters IX, X and XI deal with different aspects of this issue in the case of differentiable functions. In Chapter IX the size of the alphabet (finite) remains fixed, but the number of output vertices of the finite networks that compute the approximations is allowed to grow. The additional output vertices accommodates computing increasingly accurate approximations of the function by using more digits in a digital expansion. In Chapter X the number of output vertices remains fixed, but the size of the alphabet is allowed to grow. In Chapter XI the size of the alphabet (the alphabet is the real numbers) is fixed and the number of output vertices is also fixed, but the modules of the networks that compute the approximations are restricted to be of a specified class indexed by an integer, and the integer is allowed to grow. An example of such a class is the collection of polynomials in a fixed number of variables indexed by the degree of the polynomial. The results in Chapters IX and X give lower bounds on the computation time that are independent of the complexity of the network required to compute the function. The continuous lower bound result of Chapter IX is rarely achievable. The actual time required for the computation of a function by a network is usually much larger than the lower bound. The result in Chapter XI is a more accurate assessment because it gives conditions under which the limit of the times required for networks to compute approximations of the function is bounded below by the time required for a network to compute the function.

We consider finite approximations to a function f obtained by introducing a discrete lattice into the domain and defining the approximating function at lattice points, while the alphabet size for the networks carrying out the computations remains fixed. Even if S is a separator set for f, the lattice points in S corresponding to a given approximation to f may not be part of a separator set for the approximating function. Therefore, conditions are needed to ensure that a sequence of approximations converging to the function f yields a corresponding sequence of separator sets for the finite functions that converge in a suitable sense to the separator sets for f. In this section we study the first of two conditions that give this result. We refer to this condition as gradient-separation (or g-separation). We begin by defining approximation of a continuous function by a finite function defined on a lattice. We then study g-separation. Theorem 9.1 gives the formula for the lower bound on computing time under the condition of g-separation. In Chapter X, Theorem 10.1 gives the same lower bound formula when f is assumed to be differentiably separable (Definition 6.4).

Definition 9.1.

(i) A rectangular decomposition of R is a countable collection of half open intervals $[a_i,b_i)$ such that $R=\cup_i[a_i,b_i)$ and so that $[a_i,b_i)\cap[a_j,b_j)=\emptyset$(the empty set) unless $i=j$.

(ii) If V is a Euclidean space with standard basis $\{e_1,\ldots,e_n\}$, then a rectangular decomposition of V along the basis $\{e_i\}$ is a family of n rectangular decompositions $[a_{k\ i},b_{k\ i})$, so that