$V = \bigcup_{i(1),\ldots,i(n)} [a_{1\,i(1)}, b_{1\,i(1)})x$

$e_1[a_{2i(2)}, b_{2i(2)})e_2 \, x\ldots x[a_{n\,i(n)}, b_{n\,i(n)})e_n$.

(iii) The principal vertex of

$[a_{1\,i}, b_{1\,i})e_1 x\ldots x[a_{n\,i}, b_{n\,i})e_n$

is $(a_{1\,i}e_1 + \ldots + a_{n\,i}e_n)$.

(iv) The set of principal vertices of a rectangular decomposition is called the <u>lattice of the decomposition</u>.

If L is the lattice of a decomposition and $v \in L$, then

$B(\,v\,) = v + [a_1, b_1)e_1 x\ldots x[a_n, b_n)e_n$,

where $v = (a_1e_1 + \ldots + a_ne_n)$. The set B(v) is the <u>cube of the lattice with principal vertex v</u>. We call each $v + [a_i, b_i)e_i$ a <u>side</u> of B(v). For $x \in V$ and L the lattice of a decomposition of V, v(x) denotes that vertex such that $x \in B(\,v\,)$.

<u>Definition 9.2</u>. Suppose L_i, i=1,...,n, is the lattice of a rectangular decomposition of V_i, where V_i is a Euclidean space. Let $\epsilon > 0$ be a real number. A function

$f : \prod_{i=1}^{n} L_i \text{-->} R$

is an <u>ϵ-approximation of a function</u>

$F : V_1 x\ldots x V_n \text{--->} R$

if for each $(v_1, \ldots, v_n) \in \prod_{i=1}^{n} L_i$ and each

$(x_1,\ldots,x_n) \in B(v_1) \times \ldots \times B(v_n)$,

$|f(v_1,\ldots,v_n)-F(x_1,\ldots,x_n)|<\epsilon$.

Definition 9.3. The lattice of a rectangular decomposition of a Euclidean space is regular if the length of each side of each cube of the lattice is S, for some fixed real number S.

Definition 9.4.

(i) If R denotes the real numbers and if D is an integer greater than 1, then a radix D lattice in R is a regular lattice in R such that each vertex P of the lattice can be expressed in the form

$$\pm(a_m D^m+\ldots+a_0+\ldots+a_{-t}D^{-t})$$

where m and t are nonnegative integers and the a_i are integers between 0 and D-1. The sequence of numbers $(s,a_m,\ldots,a_{-t})$ (where s=1 if the sign of the expression is negative and s=0 otherwise) is the radix D encoding of the lattice point P.

(ii) A radix D lattice in a Euclidean space X with standard basis $\{e_1,\ldots,e_n\}$ is a regular lattice of a rectangular decomposition of X along the basis $\{e_1,\ldots,e_n\}$ in which each of the lattice points of the decomposition along each direction e_i forms a radix D lattice in Re_i.

For example, in the case D=10 the radix D encoding of a real number is the number's decimal expansion using the digits 0,...,9.

Definition 9.5. If D is an integer greater than 1 then a radix D encoding of a radix D lattice is a function which assigns to each vertex of the lattice the sequence $(s_1, \ldots, s_r)$ where each s_i is the radix D encoding of the i^{th} component of the vertex of the lattice.

In this chapter we impose the condition that a network that computes a lattice approximation of a continuous function using an alphabet {0,...,D-1} does so by computing outputs in a radix D lattice. Furthermore, the output vertices of the network carry the values a_i that are the digits of the radix D encoding of the radix D lattice. Note that the number of output vertices depends on D, and furthermore, separator sets depend on the choice and number of output vertices. If the output vertices were not explicitly determined, then this oversight would lead to a problem similar to one discussed in Chapter IV. That is, if the encoding of the lattice or if the output vertices required for the encoding of the lattice L, have not been fully specified, it is possible to hide computations by allowing the decoding to carry out some of the computations of f. In the case of a continuous function this possibility arises when one allows a network to compute an encoding of the function into a high dimensional space. This possibility already arises in the case of a finite network (as we have seen in the case of linear functions in Chapter IV) where computation time can be reduced by a judicious choice of a change of basis which amounts to expanding the number of output lines of the network. This problem shows up when one considers the size of separator sets for networks computing a function. It is, of course, true that the definition of separator set for a network is tied to the specification of the output vertices for the network and this is in turn connected with the choice of the encoding for the image points. This is illustrated by the following example.

Denote the set

{(0,0,0),(0,0,1),(0,1,0),(0,1,1),(1,0,0),(1,0,1), (1,1,0), (1,1,1)} by V. Define the function

$$f: V \times V \longrightarrow \{0,1,2,3\}$$

by the following table:

f	(0,0,0)	(0,0,1)	(0,1,0)	(0,1,1)	(1,0,0)	(1,0,1)	(1,1,0)	(1,1,1)
(0,0,0)	1	3	1	3	1	3	1	3
(0,0,1)	2	1	2	1	2	1	2	1
(0,1,0)	3	2	3	2	3	2	3	2
(0,1,1)	1	2	1	2	1	2	1	2
(1,0,0)	3	1	3	1	3	1	3	1
(1,0,1)	2	3	2	3	2	3	2	3
(1,1,0)	0	3	0	3	0	3	0	3
(1,1,1)	3	0	3	0	3	0	3	0

Table 9.1

If we consider (2,4)-networks that have a single output vertex carrying the alphabet {0,1,2,3}, then it is easy to see that the first component of V x V is a separator set for that output vertex and that the maximum separator set for the second component has two elements. Thus the minimum delay for (2,4)-networks computing this function is

$$\mathrm{INT}[\ \log_2(\ \mathrm{INT}[\ \log_4(\ 8\ )\ ]+\mathrm{INT}[\ \log_4(\ 2\ )\ ]\ )\ ]=2.$$

Suppose we allow a recoding of the set {0,1,2,3} (with the same alphabet) by the function t defined by the equations t(0)=(0,0), t(1)=(0,1), t(2)=(1,0), t(3)=(1,1)

and we consider again (2,4)-networks, but this time with the possibility of two output vertices. The table for the function now becomes Table 9.2:

f*	(0,0,0)	(0,0,1)	(0,1,0)	(0,1,1)	(1,0,0)	(1,0,1)	(1,1,0)	(1,1,1)
(0,0,0)	(0,1)	(1,1)	(0,1)	(1,1)	(0,1)	(1,1)	(0,1)	(1,1)
(0,0,1)	(1,0)	(0,1)	(1,0)	(0,1)	(1,0)	(0,1)	(1,0)	(0,1)
(0,1,0)	(1,1)	(1,0)	(1,1)	(1,0)	(1,1)	(1,0)	(1,1)	(1,0)
(0,1,1)	(0,1)	(1,0)	(0,1)	(1,0)	(0,1)	(1,0)	(0,1)	(1,0)
(1,0,0)	(1,1)	(0,1)	(1,1)	(0,1)	(1,1)	(0,1)	(1,1)	(0,1)
(1,0,1)	(1,0)	(1,1)	(1,0)	(1,1)	(1,0)	(1,1)	(1,0)	(1,1)
(1,1,0)	(0,0)	(1,1)	(0,0)	(1,1)	(0,0)	(1,1)	(0,0)	(1,1)
(1,1,1)	(1,1)	(0,0)	(1,1)	(0,0)	(1,1)	(0,0)	(1,1)	(0,0)

Table 9.2

Consider a (2,4)-network which has output vertices h_1 and h_2 where the table entries are (h_1,h_2). For example,

$f^*(\ (0,0,0),(0,0,0)\ )=$

$(h_1[\ (0,0,0),(0,0,0)\ ],h_2[\ (0,0,0),(0,0,0)\ ])=$

$(0,1)$.

A quick inspection of the table reveals now that neither h_1 nor h_2 has the first component of the space as a separator set. The separator set in VxV that has maximum cardinality in each of the components of VxV has only 3 points and this serves both h_1 and h_2. The minimal delay for a network to compute f using the encoding of Table 9.2 is

$$\text{INT}[\ \log_2(\ \text{INT}[\ \log_4(\ 3\ )\ ]+\text{INT}[\ \log_4(\ 3\ )\ ]\ )\ ] = 1.$$

We investigate the limit of the times required to compute lattice approximations of a fixed continuous function defined on some domain, where the limit is taken as the length of the sides (mesh) of the approximating lattices is decreased. We suppose that the finite networks that compute the approximating functions use a fixed finite alphabet. As the lattices are refined, i.e. as the mesh (Definition 9.7) is decreased, the number of lattice points in the domain and in the range of the function increases and therefore, at least in general, the number of output vertices for the finite networks that compute the approximations must increase. Limit results on computing time may well depend on the way the output vertices of the finite networks are specified. We have chosen one way of making a uniform designation of the output vertices that allows us to conclude that there is a close relation between the Dimension Based Lower Bound on the time required to compute an encoded version of a continuous function given in Theorem 4.2 and the Arbib and Spira lower bound for the time required for a network to compute approximations of the function. Even with this choice of encoding some restriction on the continuous functions in required. One such restriction is gradient separation (Definition 9.6).

Definition 9.6. Suppose that $F:X_1 x \ldots x X_n \text{---}> R$ is a continuously differentiable function defined on the product, X, of the Euclidean spaces $X_1, \ldots, X_n$. Suppose that for each i, U_i is a nonempty subset of X_i. Set $U=U_1 x \ldots x U_n$. Suppose that X_i has coordinates $x_{(i\ j)}$, $1 \le j \le d_i$. Let $x \in X$. For i an integer, $1 \le i \le n$, set

$$\text{grad}_{<-i>}F=$$
$$(\partial F/\partial x_{(1)}, \ldots, \partial F/\partial x_{(i-1\ d)},$$
$$\partial F/\partial x_{(i+1)}, \ldots, \partial F/\partial x_{(n\ d)}).$$

Two points, x and x′ in X_i are gradient separated by F in $U_{<-i>}$ (or g-separated by F in $U_{<-i>}$) if there exists a point z* in $U_{<-i>}$ such that

$$|\text{grad}_{<-i>}F(\ x,z^*\ )| \ne |\text{grad}_{<-i>}F(\ x',z^*\ )|.$$

Lemma 9.1. If F is a continuously differentiable function, and if x, x′ are points that are g-separated in X_i , then x and x′ are separated by F in $X_{<-i>}$.

Proof. Because x and x′ are g-separated in X_i , there is a z* in $X_{<-i>}$, such that if

$$G(\ x,x';z\ )=F(\ x,z\ )-F(\ x',z\ )$$

then, $\text{grad}_{<-i>}$ $G(x,x';z^*)\neq 0$. If $G(x,x';z^*)\neq 0$ we are finished. Suppose that

$$z^*=(z^*_{(1)},\ldots,z^*_{(i-1)},z^*_{(i+1)},\ldots,z^*_{(n)}),$$

where $z^*_{(j)}\in X_j$, and assume that $z_{(j\ k)}$, $1\le k\le d_j$, is a local coordinate system for X_j at $z^*_{(j)}$. Suppose,

$$z^*_{(i)}=(z^*_{(i\ 1)},\ldots,z^*_{(i\ d_i)}).$$

Because $(\text{grad}_{<-i>}G)[x,x';z^*]\neq 0$, it follows that $(\partial G/\partial z_{(j\ k)})[z^*]\neq 0$ for some fixed j and k. Denote by e the vector that has 1 as $(j,k)^{th}$ component and has all other components 0. Denote by L the line in the direction of e that passes through the point z*. The line L is parameterized by the function

$z^*(t)$, for $t\in R$,

$$z^*(t)=z^*+t\,e=$$

$$(z^*_{(1)},\ldots,z^*_{(j-1)},z^*_{(j\ 1)},\ldots,z^*_{(j\ k)}+t,\ z^*_{(j\ k+1)},\ldots,z^*_{(n)}).$$

Denote by G_L the restriction of G to the line L. The function G_L is a function of t and

$$\frac{d}{dt}(G_L)(0)=\left(\frac{\partial G}{\partial z_{(j\ k)}}\right)(z^*)\neq 0.$$

Because $\frac{d}{dt}(G_L)(0)\neq 0$, G_L is either increasing or decreasing near t=0. Thus for some $t^{\#}\neq 0$, $(G_L)(t^{\#})\neq 0$. Then if $z^{\#}=z^*(t^{\#})$, $G(x,x';z^{\#})\neq 0$. Therefore x and x' are separated. ∎

Lemma 9.1 shows that if F is g-separated in X_i, then the spaces X_i and (X_i/F) coincide. If F is g-separated in each X_i, then the message space for the essential revelation mechanism (cf. Defintion 6.1, Chapter VI) is the original product space $X_1 \times \ldots \times X_n$.

Computing approximations to a continuous function by means of finite networks makes it necessary to restrict the domain to be a bounded subset of Euclidean space. It is reasonable to suppose that greater accuracy of approximation requires refinement of the lattice of approximation. The next lemma justifies that supposition.

<u>Lemma 9.2</u>. Suppose that $F:K_1 \times \ldots \times K_n \text{---}{>} R$ is a continuously differentiable function, where each K_i is a compact subset (with nonempty interior) of a Euclidean space X_i of dimension d_i. Suppose that $k=(k_1,\ldots,k_n)$ is a point in the interior of $K_1 \times \ldots \times K_n$ such that (grad F)[k]$\neq 0$. Assume that for each $1 \leq i \leq n$ and $1 \leq j \leq d_i$, the elements $e_{(i\ j)}$ is the standard basis for X_i. There is then an open set U in $K_1 \times \ldots \times K_n$ and a real number $M>0$ satisfying the following conditions:

(1) $k \in U$,

(2) for some i and j, if L is a line segment in the direction $e_{(i\ j)}$ that is contained in U,

if $\epsilon>0$, and if x and x′ are elements in L such that

$$|F(x)-F(x')|<\epsilon,$$

then′

$$|x-x'|<(\epsilon/M).$$

Proof. Without loss of generality we may suppose that k=0, the origin of $X_1 x \ldots x X_n$. Denote by $x_{(i\ j)}$ the coordinate system at 0 dual to $e_{(i\ j)}$. Because (grad(F))[0]≠0, it follows that

$$\left(\frac{\partial F}{\partial x_{(i\ j)}}\right)(0)\neq 0,$$

for some i and j. Because F is continuously differentiable, there exists an open set U around 0 and a positive real number M such that for each z′ in U,

$$\left|\left(\frac{\partial F}{\partial x_{(i\ j)}}\right)(z')\right|>M.$$

Suppose that L is a line segment contained in U in the direction $e_{(i\ j)}$. Parameterize L by setting

$$z(t)=x+te_{(i\ j)}$$

for some $x\epsilon L$. Denote by F_L the composition of F with the function z(t). If $x^*\epsilon L$, then $x^*=z(t^*)$ for $t^*\epsilon R$ and

$$\left|\left(\frac{d}{dt}F_L\right)(t^*)\right|=\left|\left(\frac{\partial F}{\partial x_{(i\ j)}}\right)(z(t^*))\right|>M.$$

If x, x′ are in L, and t and t′ are such that z(t)=x and $z(t')=x'$, then the Mean Value Theorem shows that

$$|(F_L)(t)-(F_L)(t')| = |t-t'|\ |(\frac{d}{dt}F_L)(t")|$$

for some $t"\epsilon(t,t')$, where (t,t') denotes the open interval from t to t'. Therefore

$$|F(x)-F(x')|=$$

$$|t-t'|\ |(\frac{d}{dt}F_L)(t")|=|x-x'|\ |(\frac{\partial F}{\partial x_{(i\ j)}})(x")| >$$

$$|x-x'|\ M.$$

It follows that,

$$|x-x'|<|F(x)-F(x')|/M,$$

and

therefore if

$$|F(x)-F(x')|<\epsilon,$$

then

$$|x-x'|< \epsilon/M.$$ ▒

<u>Definition 9.7</u>. Suppose that L is a lattice in a Euclidean space X. The <u>mesh</u> of the lattice L is the maximum (if it exists) of the distance between adjacent vertices (along sides) of L.

<u>Definition 9.8</u>. If X is a Euclidean space with standard basis $\{e_1,\ldots,e_n\}$ and if L and L' are lattices of rectangular decompositions of X along the basis $\{e_1\ldots,e_n\}$, then we say that L' is a <u>refinement</u> of L if each vertex of L is a vertex of L'.

Theorem 9.1 relates the Dimension Based lower bound for a gradient-separable function $F:X_1 x \ldots x X_n \text{--->} R$ to the lower bound given by Arbib and Spira for ϵ-approximations to F. The relationship between these two lower bounds is established in two steps. The first step is to show that the function F(x) can be replaced by a function defined only at lattice points and that takes values of the form $\Sigma\ a_p(x)D^p$, where D is a positive integer. The second step is to show that if v, $v' \epsilon X_i$ with $v \neq v'$, and if F is ϵ-approximated by a function f that also takes values of the form $\Sigma\ a_p\ D^p$, one can choose a sufficiently small integer p and a point $z \epsilon X_{-i}$ such that the coefficient of D^p in the radix D encoding of f($v \int_i z$) is different from the coefficient of D^p in the radix D encoding of f($v' \int_i z$). Both of these steps are carried out in Lemma 9.4. The proof of the second step is a tedious argument that uses linear approximations of F($v \int_i z$) and F($v' \int_i z$). The next lemma, Lemma 9.3, is used in the proof of the second step to establish that if f($v \int_i z$) and f($v' \int_i z$) have the same coefficient of D^p in their radix D expansions, then a small change in z to a point z' changes the values of f($v \int_i z'$) and f($v' \int_i z'$) so that the coefficients of D^p are different. Lemma 9.3 refers to the relationships shown in Figure 9.1.

The argument is carried by references to that diagram.

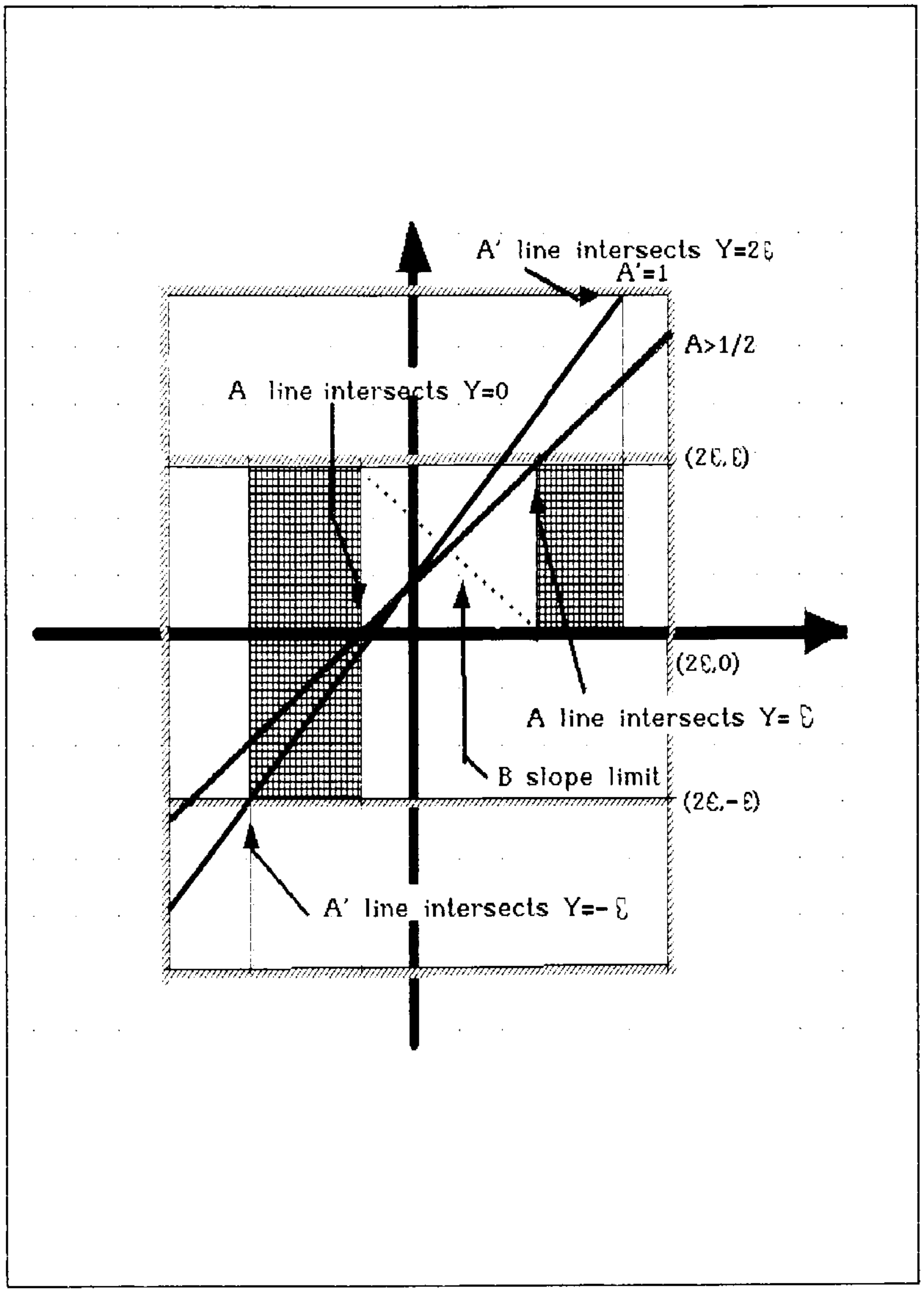
A' line intersects Y=2ε
A'=1
A>1/2
A line intersects Y=0
(2ε, ε)
(2ε,0)
A line intersects Y= ε
B slope limit
(2ε,-ε)
A' line intersects Y=- ε

Figure 9.1

Lemma 9.3. Suppose that A, B, A', B', α, β, ϵ are real numbers so that $0<A<A'$, $2A>A'$, $0\leq\alpha$, $\beta\leq\epsilon$. If $0\leq B<B'<A<A'$, then there is a nonempty open interval I contained in the interval $(-2\epsilon/A, 2\epsilon/A)$ that satisfies one of the following conditions:

(i) for each $t\in I$, $\epsilon<At+\alpha$, $A't+\alpha<2\epsilon$,

$0\leq B't+\beta<Bt+\beta<\epsilon$;

(ii) for each $t\in I$, $At+\alpha<0$, $-\epsilon<A't+\alpha$,

$0\leq B't+\beta<Bt+\beta<\epsilon$.

If $B'<B<0<A<A'$ and $|B'|<A'$, then there is a nonempty open interval I contained in the interval $(-2\epsilon/A, 2\epsilon/A)$ so that one of the following conditions is satisfied:

(iii) $At+\alpha<0$, $-\epsilon<A't+\alpha$, $0\leq B't+\beta<Bt+\beta<\epsilon$;

(iv) $At+\alpha>\epsilon$, $A't+\alpha<2\epsilon$, $0\leq B't+\beta<Bt+\beta<\epsilon$.

Proof. If we divide the inequalities in each of the conditions (i)-(iv) by A', and if we replace ϵ by $\epsilon'=\epsilon/A'$, we can assume that $A'=1$ and that $A>1/2$. The condition that $I\subseteq(-2\epsilon/A,2\epsilon/A)$ will be satisfied if we can choose an interval $I'\subseteq(-2\epsilon',2\epsilon')$ so that the following translations of (i)-(ii) are satisfied:

(i') for each $t\in I$, $\epsilon< At+\alpha$, $t+\alpha<2\epsilon'$, $0\leq B't+\beta<\epsilon'$;

(ii') for each $t\in I$, $At+\alpha<0$, $-\epsilon'<t+\alpha$, $0\leq B't+\beta<\epsilon'$,

or that the following translations (iii') and (iv') of conditions (iii) or (iv) are satisfied:

(iii') $At+\alpha<0$, $-\epsilon'<A't+\alpha$, $0\leq Bt+\beta$, $B't+\beta<\epsilon'$;

(iv′) $At+\alpha>\epsilon'$, $A't+\alpha<2\epsilon'$, $0\leq B't+\beta<\epsilon'$.

Now refer to Figure 9.1. In Figure 9.1 a line is labelled with its slope. The vertical axis is Y and the horizontal axis is the t axis. In order that one of the pairs of conditions

$$At+\alpha>\epsilon',\ t+\alpha<2\epsilon'$$

or

$$At+\alpha<0,\ t+\alpha>-\epsilon$$

be satisfied for some values of t, it is both necessary and sufficient that values of t can be chosen from the section of the t-axis that occurs in the "hatched" regions shown in Figure 9.1. Denote by K the open interval along the t-axis between the points where the line $Y=B't+\beta$ intersects the line $Y=0$ and the line $Y=\epsilon$. In order that the conditions

$$0\leq B't+\beta<B't+\beta<\epsilon,\ \text{for } B'>0$$

or

$$\epsilon>B't+\beta>0,\ \text{for } B'<0,$$

to be satisfied for values of t on an interval J, it suffices that there are values of t on the interval J that lie in K. Therefore, to prove that an interval I exists that satisfies the conditions (i′) or (ii′), when $B\geq 0$, it suffices to show that K intersects the interior of the hatched areas in Figure 9.1. Equivalently, it will suffice to show that the section of the line with equation $Y=B't+\beta$ that lies between the lines $Y=0$ and $Y=\epsilon$ intersects the interior of the hatched areas in Figure 9.1. If $B'>0$, the line $B't+\beta$ intersects the hatched area in interior points because $B'<A$. Indeed, the length of the intersection of the hatched areas with the t-axis is $(3\epsilon-\epsilon/A)$, the length of K is ϵ/B', and the length of the segment of the t-axis between the value of t where the $t+\alpha$ line intersects the line $Y=-\epsilon$ and where the line $Y=t+\alpha$ intersects the line $Y=2\epsilon$ is 3ϵ. Thus the length of the intersection of the hatched section of the t-axis and K is at least $(3\epsilon-\epsilon/A)+(\epsilon/B')-3\epsilon=\epsilon(1/B'-1/A)>0$. The case when $B'<0$ is handled in a similar fashion.▒

Lemma 9.4 is

Lemma 9.4. Assume that $X_1,\ldots,X_n$ are Euclidean spaces of dimensions $d_1,\ldots,d_n$, respectively. Assume that K_j is a compact subset of X_j with nonempty interior K_j^0. Set $K=K_1 x\ldots x K_n$, set $K^0=K_1^0 x\ldots x K_n^0$ and assume that $F:K\text{-->}R$ (R the real numbers) is a positive and continuously differentiable function. Assume that $D>1$ is an integer and:

(1) for each positive integer m, and each $1\le j\le n$, $L_j(m)$ is a regular radix D lattice in X_j of mesh D^{-m};

(2) if m and m′ are positive integers and $m\ge m'$, then $L_j(m)$ is a refinement of $L_j(m')$;

(3) if $L(m)=L_1(m)x...xL_n(m)$, then for each integer p there is a function $a_p(x;m)$ defined on $K\cap L(m)$ with values in $\{0,...,D-1\}$ so that the $a_p(x;m)$ satisfy the following conditions;

(i) for each m, there is an integer $A(m)>0$ so that $a_p(x;m)=0$ if $p<-A(m)$;

(ii) if $m'\geq m$, then $A(m')\geq A(m)$ and $\lim_{j\to\infty} 1/A(j)=0$;

(iii) if $m'\geq m$, then for each $p\geq -A(m)$ and each x in $L(m)\cap K^0$, $a_p(x;m)=a_p(x;m')$;

(iv) if $\epsilon>0$ is a real number, then there is an integer $M(\epsilon)$ so that for $M>M(\epsilon)$, the function $f^{(M)}(x)=\Sigma_p\, a_p(x;M)D^p$ is an ϵ-approximation of F on K;

(4) for some integer m*, and a fixed integer i, $1\leq i\leq n$, there are vertices v and v′ in $L_p(m^*)\cap K_i$ that are gradient separated in $K_{<-i>}{}^0$.

Then there is an integer $J(m^*)>0$, and for each $p>J(m^*)$ an integer $M(p)>m^*$ such that if $m>M(p)$ the vertices v and v′ are separated by $a_{-p}(x;m)$ in

$L_{<-j>}(m)$.

Proof. The proof breaks into three sections. In the first section, we show that the function F(x) can be replaced by a function f(x) that is defined only on points x that are in $K \cap \cup_m L(m)$ and that has as values series of the form $\Sigma_p\ a_p(x)D^p$ where the $a_p(x)$ are functions of x. This shows that it is possible to talk, unambiguously, about the coefficient of D^p in the radix D representation of F(x) as long as x is chosen from the union of lattices $\bigcup_m L(m)$. In the second section we establish that for a fixed i, if F is gradient-separated by $K_{<-i>}$, and if v, $v' \in X_i$, $v \neq v'$, are the vertices given in condition (iv), then it is possible to choose linear approximations of F($v\int_i x$) and F($v'\int_i x$), $x \in X_{<-i>}$, so that the directional derivatives of F($v\int_i x$) and F($v'\int_i x$) along some coordinate direction satisfy the conditions placed on A and B in Lemma 9.3. The last part of the proof uses the linear approximations to argue that if p is a sufficiently small integer, and if for a $y_0 \in X_{<-i>}$

$$grad_{<-i>}F(v\int_i y) \neq \pm grad_{<-i>}F(v'\int_i y),$$

and if

$$a_p(v\int_i y_0)=a_p(v'\int_i y_0),$$

then it is possible to choose a small change in y to a value y_1 such that either

$$a_p(v\int_i y_0)\neq a_p(v\int_i y_1)$$

and

$$a_p(v'\int_i y_0)=a_p(v'\int_i y_1)$$

or

$$a_p(v'\int_i y_0)\neq a_p(v'\int_i y_1)$$

and

$$a_p(v\int_i y_0)=a_p(v\int_i y_1).$$

Set $X=\prod_j X_j$. Suppose that the lattices $L_j(m)$ are the lattices of a rectangular decomposition of X_j along the standard basis $\{e_{(j\ k)}\}$, where for a fixed j the vectors $\{e_{(j\ k)}\}$ form a basis for X_j . Set

$$L=\bigcup_{m'} L(m').$$

Note that L is dense in X and the vertices of L are also dense along each line that passes through a vertex of L and is parallel to one of the basis elements $e_{(j\ k)}$. Condition 3(iii) guarantees that if $x\in L(m)\cap K^0$, then for $p\geq -A(m)$ and $m'\geq m$, $a_p(x;m')$ is independent of m'. In particular, this shows that for each $x\in L\cap K^0$, $x\in L(m')\cap K^0$ for m' sufficiently large. Furthermore, for each p, $a_p(x;m'')$ is independent of m'', if m'' is sufficiently large. Set

$$\lim_{m'\to\infty} a_p(x;m')=a_p(x).$$

There is an integer P, such that for each $x\in L\cap K^0$ and each $p>P$, $a_p(x)=0$. To see this, note first that $|F|$ is bounded on the compact set K. Furthermore, condition 3 (iv) assures that for $\epsilon=1$ there is a real number M(1) such that if M>M(1), then

$$f^{(M)}(x)=\Sigma_p\, a_p(x;M)D^p$$

is a 1-approximation of F on K. Therefore, for each $x \in L(M)\cap K$,

$$|f^{(M)}(x)-F(x)|<1$$

It follows that $|f^{(M)}(x)|$ is bounded on K by $1+\max_{x\in K}|F(x)|$. Therefore, $a_p(x;M)=0$ for p sufficiently large and $x\in L\cap K$. The series

$$\Sigma_p\, a_p(x)D^p$$

converges for all $x\in L\cap K$. This is because $0\leq a_p(x)<D$, and for p sufficiently large $a_p(x)=0$. Set

$$f(x)=\Sigma_p\, a_p(x)D^p\ .$$

We next show that for each $x\in L\cap K^0$,

$$F(x)=f(x).$$

If $x\in L\cap K^0$, then there is an integer N such that if $m>N$, $x\in L(m)\cap K^0$. For each such m it follows from condition 3 (i) that there is an integer A(m) such that $a_p(x;m)=0$ if $p<-A(m)$. Furthermore, condition 3(iii) implies that for $x\in L(m)\cap K^0$ and $p\geq -A(m)$,

$$\lim_{m'\to\infty} a_p(x;m')=a_p(x;m)=a_p(x).$$

Therefore

$$f(x)=\Sigma_{p\geq -A(m)}a_p(x;m)D^p+\Sigma_{p<-A(m)}a_p(x)D^p$$

and therefore

$$|f(x)-f^{(m)}(x)|=$$

$$|\Sigma_{p<-A(m)}a_p(x)D^p|\leq(D-1)D^{-A(m)}[\Sigma_{p\leq 0}D^p]=$$

$$D^{-A(m)+1}\ .$$

If $\epsilon>0$ is a real number, condition 3 (iv) states that there is an integer $M(\epsilon/2)$ so that for each $M>M(\epsilon/2)$, the function $f^{(M)}(x)$ is an $\epsilon/2$-approximation of F on K. Choose M^* so large that for $m>M^*$ and $x\in L(m)\cap K^0$, $f^{(m)}$ is an $\epsilon/2$-approximation for F on K and $D^{-A(m)+1}<\epsilon/2$. Then

$$|f(x)-F(x)|=$$

$$|f(x)-f^{(m)}(x)+f^{(m)}(x)-F(x)|\leq\epsilon.$$

We have established that for $x\in L\cap K^0$, and for $\epsilon>0$ a real number, $|f(x)-F(x)|\leq\epsilon$. Therefore,

$$f(x)=F(x).$$

It follows that the function f(x) determines uniquely the radix D representation of F(x) for each $x\in L\cap K^0$. Furthermore, for each integer p and each $x\in L\cap K^0$, $a_p(x)$ is a function of x. This completes the first section of the proof.

We turn to the second section where we construct the linear approximations required. For the integer i and the vertices v and v′ fixed in condition (4), and for each $z\in K_{<-i>}$, set

$$g(z)=F(v\textstyle\int_i z)$$

and

$$g'(z)=F(v'\textstyle\int_i z).$$

If x and y are elements of K, then denote by $x\cdot y$ the dot product of x and y (that is, the inner product determined by the basis $\{e_{(j\ k)}\}$). Because v and v′ are gradient-separated in $K_{<-i>}{}^{0}$, there is a point $z^*\in K_{<-i>}{}^{0}$ such that

$$|(\mathrm{grad}_{<-i>}g)[\ z^*\ ]|\neq|(\mathrm{grad}_{<-i>}g')[\ z^*\ ]|.$$

Therefore, there are integers a* and b* such that

$$(\mathrm{grad}_{<-i>}g)[\ z^*\ ]\cdot e_{(a^*\ b^*)}\neq$$
$$\pm(\mathrm{grad}_{<-i>}g')[\ z^*\ ]\cdot e_{(a^*\ b^*)}.$$

Because g and g′ are continuously differentiable on K, there is a ball S in $K_{<-i>}{}^{0}$ such that $z^*\in S$ and such that for each $x\in S$,

$$(\mathrm{grad}_{<-i>}g)[\ x\ ]\cdot e_{(a^*\ b^*)}\neq$$
$$\pm(\mathrm{grad}_{<-i>}g')[\ x\ ]\cdot e_{(a^*\ b^*)}\ .$$

For a sufficiently large M′, if M≥M′ then L(M)∩S is nonempty. Choose a w∈L(M′)∩S. Then w∈L(M)∩S for all M≥M′. Denote by q the function from R to $K_{<-i>}$ given by the equation

$$q(\ t\ )=w+te_{(a^*\ b^*)}\ .$$

Because we can interchange g and g′ and reverse the direction of $e_{(a^*\ b^*)}$, if necessary, we can assume that if

$$B^*=(\mathrm{grad}_{<-i>}g')[\ w\ ]\cdot e_{(a^*\ b^*)}\ ,$$

then

$$0\leq|B^*|<(\mathrm{grad}_{<-i>}g)[\ w\ ]\cdot e_{(a^*\ b^*)}=A^*\ .$$

Because A*>0, for |t| sufficiently small,

$$(\mathrm{grad}_{<-i>}g)[\ w+te_{(a^*\ b^*)}\ ]\cdot e_{(a^*\ b^*)}>0.$$

If B*=0, then either

$$(\mathrm{grad}_{<-i>}g')[\ w+te_{(a*\ b*)}\ ]\cdot e_{(a*\ b*)} = 0$$

for t in a neighborhood of 0 where

$$(\mathrm{grad}_{<-i>}g)(w+te_{(a*\ b*)}\cdot e_{(a*\ b*)} > 0$$

or we can replace w

with a $w'=w+t'e_{(a*\ b*)}$ such that

$$|B'*|=|(\mathrm{grad}_{<-i>}g')(\ w\ )\cdot e_{(a*\ b*)}|>0,$$

and

$$(\mathrm{grad}_{<-i>}g)(\ w\ )\cdot e_{(a*\ b*)}=A'*>0.$$

Therefore, we assume that either $0<|B*|<A*$ or that $g'(\ t\ )$ is constant along the line parameterized by $q(\ t\ )$ near w.

Set

$$G(\ t\ )=g(\ q(\ t\ )\ )$$

and

$$G'(\ t\ )=g'(\ q(\ t\ )\ ).$$

By definition, the values $G(\ 0\ )$ and $G'(\ 0\ )$ are the values $F(\ v\int_i w\ )$ and $F(\ v'\int_i w\ )$. Since F and f are the same function on $L \cap K$,

$$G(\ 0\ )=f(\ v\int_i w\ )$$

and

$$G'(\ 0\ )=f(\ v'\int_i w\ ).$$

Thus

$$G(\ 0\ )=\Sigma_p\ a_p(\ v\int_i w\ )D^p$$

and

$$G'(\ 0\ )=\Sigma_p\ a_p(\ v'\int_i w\ )D^p.$$

For an arbitrary integer T, if we write

$$a'_T=\Sigma_{p>T}\, a_p(\, v\int_i w\,)D^p\ ,$$

and

$$b_T=\Sigma_{p>T}\, a_p(\, v'\int_i w\,)D^p$$

then we can write

$$G(\,0)=a'_T+a_T(\,v\int_i w\,)D^T+\alpha_T$$

and

$$G'(\,0\,)=b_T+a_T(\,v'\int_i w\,)D^T+\beta_T$$

where $0\le\alpha_T,\ \beta_T\le D^T$. Because $G(\,t\,)$ and $G'(\,t\,)$ are differentiable, there are functions $\mu(\,t\,)$ and $\mu'(\,t\,)$, such that

$$G(\,t\,)=A^*\,t+\mu(\,t\,)+G(\,0\,),$$

$$G'(\,t\,)=B^*\,t+\mu'(\,t\,)+G'(\,0\,)$$

and such that

$$\lim_{t\to 0}\mu(\,t\,)/t=\lim_{t\to 0}\mu'(\,t\,)/t=0.$$

The approximations above require bounds that conform to the hypothesis of Lemma 9.3. To achieve this we consider two cases that depend on the relation of B* to 0.

First suppose that $0<B^*<A^*$. We can choose an interval U around 0 so that for each $t\in U$,

$$|\mu(\,t\,)/t|<\min(\,A^*/4,(A^*-B^*)/4\,)=C$$

and

$$|\mu'(\,t\,)/t|<\min(\,B^*/2,(A^*-B^*)/2)\,).$$

Set

$$B=B^*/2,\ B'=B^*+(A^*-B^*)/2=(A^*+B^*)/2,\ A=A^*-C,$$

and

$$A'=A^*+A^*/4=5A^*/4.$$

Set

$$A^{\#}(t)=A^*+\mu(t)/t$$

and set

$$B^{\#}(t)=B^*+\mu'(t)/t.$$

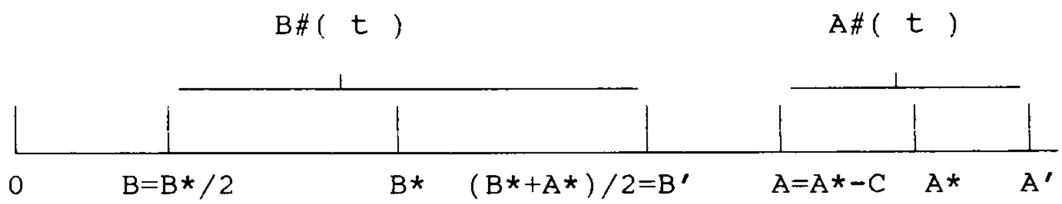

Diagram 9.1

Clearly (see Diagram 9.1),

$$0<B<B^*<B'<A<A^*<A'$$

and

$$2A>A'.$$

Also

$$B^{\#}(t)=B^*+\mu'(t)/t>B^*-B^*/2=B$$

and

$$B^{\#}(t)<B^*+(A^*-B^*)/2=B'.$$

Further

$$A^{\#}(t)=A^*+\mu(t)/t>A^*-\min((A^*-B^*)/4,A^*/4)=A$$

while

$$A^{\#}(t)<A^*+A^*/4=A'.$$

Finally

$$2A=2(A^*-C)=6A^*/4>5A^*/4=A'.$$

It follows that

$$0<B<B^{\#}(t)<B'<A<A^{\#}(t)<A',$$

and

$$2A>A'.$$

Second, suppose that $B^*<0$. Set $C^*=\min(A^*-|B^*|, |B^*|)$.

Choose an interval U so that for $t \in U$,

$$|\mu(t)/t|<C^*/4$$

and such that

$$|\mu'(t)/t|<A^*/4.$$

Then set

$$B'= B^*-C^*/4, B=B^*+C^*/4, A=3A^*/4, \text{ and } A'=5A^*/4.$$

Then (See Diagram 9.2)

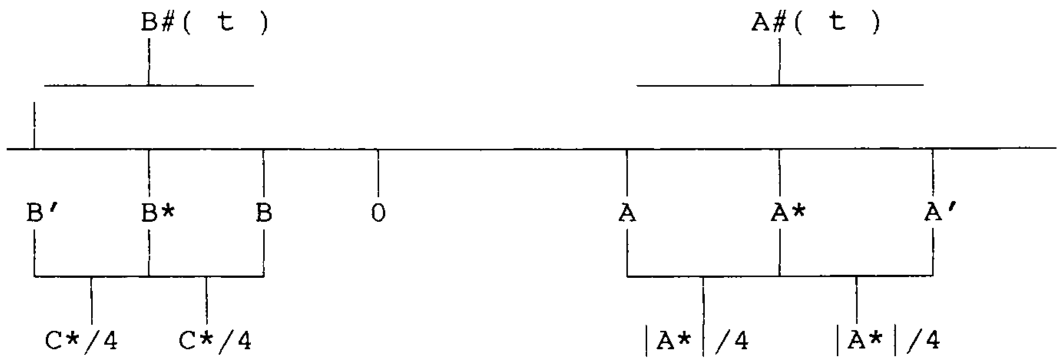

Diagram 9.2

$$B' < B^{\#}(t) < B < 0 < A < A^{\#}(t) < A'$$

while

$$2A > A' \text{ and}$$

$$|B'| < A'.$$

We now begin the last section of the proof, making use of the linear approximations. For $t \in U$ and T a nonnegative integer, set

$$G(t) = A^{\#}(t)t + a_T + a_T(v\textstyle\int_i w)D^T + \alpha_T$$

and set

$$G'(t) = B^{\#}(t)t + b_T + a_T(v'\textstyle\int_i w)D^T + \beta_T .$$

Choose T so small (i.e. $|T|$ so large) that the interval

$$U'(T) = (-2D^T/A,\ 2D^T/A)$$

is contained in U. If $q(t) \in K$, then

$$G(t) = f(v\textstyle\int_i q(t)) = \Sigma_p\, a_p(v\textstyle\int_i q(t))D^p$$

and

$$G'(t) = \Sigma_p\, a_p(v'\textstyle\int_i q(t))D^p .$$

Set

$$a_p(v\textstyle\int_i q(t)) = a^{\#}_p(t)$$

and set

$$a_p(v'\int_i q(t))=b^{\#}_p(t).$$

To complete the proof of the lemma it will suffice to show that for each $-p<T$, there is a $t\in U'$ so that $a^{\#}_{-p}(t)\neq b_{-p}^{\#}(t)$. Because

$$G(t)=A^{\#}(t)t+a^{\#}_{-p}(0)D^{-p} +\alpha_{-p} +a_{-p}$$

and

$$G'(t)=B^{\#}(t)t+b^{\#}_{-p}(0)D^{-p}+\beta_{-p}+b_{-p} ,$$

it will suffice to show that if

$$a_{-p}(0)=b_{-p}(0),$$

then there is a $t\in U'$ so that one of the following conditions is satisfied:

(i) $a^{\#}_{-p}(0)<D-1$ and $a^{\#}_{-p}(t)=a^{\#}_{-p}(0)+1$,

$b^{\#}_{-p}(t)=b^{\#}_{-p} (0)$,

or

$a^{\#}_{-p}(0)=b^{\#}_{-p}(0)=D-1$,

$a^{\#}_{-p}(t)=0$, and $b^{\#}_{-p}=D-1$;

(ii) $0<a^{\#}_{-p}(0)$ and $a^{\#}_{-p}(t)=a^{\#}_{-p}(0)-1$,

$b^{\#}_{-p}(t)=b^{\#}_{-p}(0)$,

or

$a^{\#}_{-p}(0)=b^{\#}_{-p}(0)=0$, and $a^{\#}_{-p}(t)=D-1$,

$b^{\#}_{-p}(t)=0$.

Condition (i) is implied by the conditions:

(I) $a^{\#}_{-p}(0)<D-1$, $D^{-p}<A^{\#}(t)t+\alpha_{-p} <2D^{-p}$

$D^{-p}>B^{\#}(t)t+\beta_{-p} \geq 0$,

while condition (ii) is implied by the

condition,

$$(\text{II})\quad 0\le a^{\#}_{-p}(\ 0\ ),\quad -D^{-p}\le A^{\#}(\ t\ )t+\alpha_{-p}<0,$$

$$D^{-p}>B^{\#}(\ t\ )t+\beta_{-p}>0\ .$$

Because $A\le A^{\#}(\ t\ )\le A'$ and $B\le B^{\#}\le B'$, if $B\ge 0$,

$$At+\alpha_{-p}\le A^{\#}(\ t\ )t+\alpha_{-p}\le A't+\alpha_{-p}\ .$$

Therefore (I) is satisfied if ;

$$(\text{I}')\quad D^{-p}<At+\alpha_{-p}\ ,\ A't+\alpha_{-p}<2D^{-p}\ \text{ and}$$

$$0<\ Bt+\beta_{-p}<D^{-p}.$$

Similarly, condition (II) is satisfied if;

$$(\text{II}')\quad At+\alpha_{-p}<0,\ -D^{-p}<A't+\alpha_{-p}\ ,\ 0\le B't+\beta_{-p}<D^{-p}\ ,$$

Lemma 9.3 shows that there is an interval I contained in the interval $(-2D^{-p}/A, 2D^{-p}/A)$ so that for $t\in I$, either (I′) or (II′) is satisfied. If m is sufficiently large, then there are vertices of L(m) that lie in the interval I. Therefore the function $a_{-p}(\ x;m\ )$ separates v and v′.▒

<u>Theorem 9.1</u> . Assume the hypotheses (1)-(4) of Lemma 9.4. Assume that each pair of points x and x′ in K_i are g-separated by F. Then for ϵ sufficiently small and M sufficiently large, if the network $C(\epsilon,M)$ computes $f^{(M)}(\ x\ )$ in time T, then

$$T\ge \mathrm{INT}[\ \log_D(\ \Sigma_i \dim X_i\ )\ ].$$

Proof. For M sufficiently large, we can assume that for each $1\le j\le n$, all the vertices of a cube $Q(j)$ of $L_j(M)$ are contained in the interior of K_j . The lattice $L_j(M)$ has each side of $Q(j)$ split into D points. The vertices of $L_j(M)$ that are contained in $Q(j)$ consists of $D^{d(j)}$ points where $d(j)=\dim X_j$. Lemma 9.4 shows that if S is sufficiently large, then each pair of vertices in $Q(j)$ form a separator set for the function $a_{-S}(x)$. Therefore, the vertices form a separator set for the function $f^{(M)}(x)$. The lower bound given by Arbib and Spira for finite functions(c.f. Chapter II, Theorem 2.2) shows that the minimum computing time for $f^{(M)}(x)$ is at least

$$\text{INT}\ [\log_D(D^{d(j)})]=d(j)=\dim X_j\ .\ ▒$$

Chapter X

Separator Sets for Smooth Functions II; Differentiable Separability

In this chapter, as in the Chapter IX, we relate the computation lower bound based on dimension for a network that computes a real valued function F to a lower bound on the computing time required for networks that compute ϵ-approximations of F. A principal distinction between the approach of this chapter and that of Chapter IX is the type of ϵ-approximation used. In Chapter IX the networks that compute the ϵ-approximation use a fixed finite alphabet, and compute a radix encoding of the values of the function F. In Chapter IX, as ϵ decreases in size, the number of output vertices of the networks carrying out the computation increases. In this chapter, the network used to compute the approximation has one output vertex, but the alphabet used by the network grows in cardinality as ϵ decreases in size.

The rest of this chapter is organized as follows. In section 10.1 we introduce separator functions and discuss their uses. Separator functions are a convenient method of constructing the separator sets introduced in Chapter III. Separator sets for a function $F:X_1x...xX_n$--->R are used to establish a lower bound on the time required to compute F by networks whose modules are continuous functions. Recall that a locally Euclidean subspace S_i of X_i is a separator set if for each pair of points s, s′ of S_i, with s≠s′, there is a point $w_i(s,s') \in X_{<-i>}$, such that

$$F(s \int_i w_i(s,s')) \neq F(s' \int_i w_i(s,s')).$$

We make the natural assumption that the relation w_i is a function of s and s′. If w_i is a function then we use the notation W_i. However, this is not by itself adequate to ensure that an ϵ-approximation to F has sufficiently many (lattice) points in a separator set. A mild additional condition is imposed on the function W_i to ensure that separator sets for the function F have in them a collection of lattice points that form separator sets for functions that ϵ-approximate F. Functions that satisfy this additional condition are separator functions.

In order that a lattice function f be an ϵ-approximation of a function F, the mesh of the lattice on which the approximation is defined must be sufficiently small. Section 10.2 analyzes a relation between ϵ and the mesh of the lattice used for the approximation. The section ends with a theorem stating that if $F:X_1 x \ldots x X_n \text{--->} R$ is a function with locally Euclidean separator sets $S_i \subseteq X_i$ and separator functions, then there is a lower bound on the computing time for ϵ-approximations that can be stated in terms of the dimensions of the S_i when the mesh of the lattice and ϵ are nicely related.

In Section 10.3 we show that if F is differentiably separated of rank $(r_1,...,r_n)$, then there is a separator submanifold S_i of dimension r_i in X_i and a separator function W_i associated to S_i. The section ends with the relation between the Dimension Based lower bound on the time required to compute a function F and the time required for finite networks to compute lattice approximations of the function F.

Section 10.1

We begin with some notation.

<u>Notation</u>: If Y is a set, then

$$\text{diag}(Y)=\{(y,y) \in YxY\}.$$

<u>Definition 10.1.</u> Assume that for $1 \le i \le n$, X_i is a Euclidean space and suppose that S_i is a subset of X_i. Let U_i be a nonempty neighborhood of the origin in X_i. A function $F:X_1x...xX_n$--->R is said to have W_i as a separator function in an open set $U_1x...xU_n$ along the set S_i if

$$W_i:(U_i \cap S_i)x(U_i \cap S_i)-\text{diag}(U_i \cap S_i) \longrightarrow S_{<-i>}$$

and there is a positive number M such that the following condition is satisfied:

For each

$$(y,y') \in (U_i \cap S_i) x (U_i \cap S_i) - \mathrm{diag}(U_i \cap S_i) \ ,$$

$$|F(y,W_i(y,y'))-F(y',W_i(y,y'))| \geq M \ |y-y'|.$$

Note that if F has a separator function in a neighborhood $U_1 x \ldots x U_n$ of a point $p=(p_1,\ldots,p_n)$ along a set S_i, then the function F also has a separator function on a neighborhood $V_i \int_i U_{<-i>}$ of p if $p_i \in V_i < U_i$.

We can extend this definition to a function F defined on a product of differentiable manifolds. We use the definitions and conventions found in [6].

<u>Definition 10.2.</u> Suppose that $X=\prod_{i=1}^{n} X_i$ is a product of differentiable manifolds X_i, dim $X_i=d(i)$, and suppose that F:X--->R is a real valued differentiable function. Suppose that $p=(p_1,\ldots,p_n)$ is a point of X and assume that for each i there is a coordinate neighborhood V_i based at p_i with coordinate functions $\{\varphi_{i\ 1},\ldots,\varphi_{i\ d(i)}\}$ that map V_i into the Euclidean space E_i. Suppose that for each $1 \leq i \leq n$, S_i is a subset of X_i. For each $y \in V_i$, set $\varphi_i(y)=(\varphi_{i\ 1}(y),\ldots,\varphi_{i\ d(i)}(y))$. Denote by U_i the open set in E_i that is the image of V_i under the function φ_i. The function F is said to have separator

function W_i in the coordinate neighborhood $V_1 x \ldots x V_n$ of the point $p=(p_1,\ldots,p_n)$ along the set S_i in the coordinates $(\varphi_1,\ldots,\varphi_n)$ if there are open sets V'_i, $p_i \in V'_i \subseteq V_i$, and a function

$$W_i:(V'_i \cap S_i) x (V'_i \cap S_i) - \mathrm{diag}(V'_i \cap S_i)$$

such that the function

$(\prod_{j\neq i} \varphi_j)\cdot W_i \cdot \varphi_i^{-1}$ is a separator function for $F\cdot(\prod\varphi_i)^{-1}$ in the open set $\prod_j U_j$ along the sets $\{\varphi_i(S_i)\}$(c.f. Figure 10.1).

$$\begin{array}{ccc}
V'_i \cap S_i & \xrightarrow{W_i} & V_1\cap S_1 \; x \ldots x \; V_i \cap S_i \; x\ldots x \; V_n \cap S_n \\
\downarrow & & \downarrow \prod \varphi_j \\
\varphi(V'_i)\cap\varphi_i(S_i) & \xrightarrow[(\prod\varphi_i)^{-1}\cdot W_i\cdot\varphi_i]{} & U_1\cap\varphi_1(S_1) \; x \ldots x \; U_i\cap\varphi_i(S_i) \; x\ldots x \; U_n\cap\varphi(S_n)
\end{array}$$

Figure 10.1.

In the following lemma, we show that the property of having a separator function is coordinate free.

Lemma 10.1. Suppose that for each $1 \le i \le n$, X_i is a C^1-differentiable manifold, and assume that S_i is a submanifold of X_i. Assume that $p=(p_1,\ldots,p_n)$ is a point of $X_1 x \ldots x X_n$ and suppose that $U_1 x \ldots x U_n$ is a coordinate neighborhood of p with two sets of coordinates $(\varphi_1,\ldots,\varphi_n)$ and $(\Theta_1,\ldots,\Theta_n)$ defined on $U_1 x \ldots x U_n$. If F has a separator functions W_i along the set S_i in the coordinates $(\varphi_1,\ldots,\varphi_n)$, then F has a separator function along the set S_i in the coordinates $(\Theta_1,\ldots,\Theta_n)$.

Proof. We may assume, without loss of generality, that the X_i are Euclidean spaces and that the coordinates $(\varphi_1,\ldots,\varphi_n)$ are linear coordinates on the X_i relative to the standard basis. The functions $\varphi_{i\ j}$ (where $\varphi_i=(\varphi_{i\ 1},\ldots,\varphi_{i\ d(i)})$ are functions with continuous derivatives on a compact subset of X_i that contains U_i. It follows immediately from the Mean Value Theorem, by summing the inequalities for the components of the φ_i, that there is a real number N which is independent of i, such that if p, $p' \in U_i$, then

$$|\varphi_j(\ p\ )-\varphi_j(\ p'\ )| \ge M|\Theta_j(\ p\ )-\Theta_j(\ p'\ )|.$$

Suppose that W_i is a separator function for F in the coordinates φ_i. Then for $p, p' \in U_i \cap S_i$, $p \neq p'$,

$$|F(p,W_i(p,p'))-F(p',W_i(p,p'))| \geq$$

$$M|\varphi_j(p)-\varphi_j(p')| \geq N|\Theta_j(p)-\Theta_j(p')|. ▒$$

Section 10.2

Recall that if a map $g:X \text{---}> Y$ is a submersion, then the Jacobian Jg of the mapping g has rank equal to dim(Y) at each point of X. If dim(X)-dim(Y)>0, and if the map g is a submersion, then it is known(c.f. [6, p.9]) that the map can be linearized. 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 $q \in U$, $g(q)=(x_1(q),\ldots,x_m(q))$.

In the following theorem, we show that if a function is differentiably separable at a point on a C^3-manifold, then the function has separator functions in a neighborhood of the point.

<u>Theorem 10.1.</u> Suppose that for $1 \leq i \leq n$, X_i is a C^3-manifold of dimension d(i) and suppose that $F:\prod_1^n X_i \text{---}> R$ is a C^3-function. If $p=(p_1,\ldots,p_n) \in \prod X_i$ and if F is differentiably separable at p, then for each $1 \leq i \leq n$, F has a separator function W_i on a neighborhood $U_1 x \ldots x U_n$ of the point p.

The proof of this theorem is intricate. Before giving the general proof, the argument is given in the context of a simple example.

<u>Example.</u> Suppose that U_1 and U_2 are open balls of radius R>0 around the origin of R^2. Choose R>r>0. Assume that U_1 has coordinates (x,y) and assume that U_2 has coordinates (z,w). Let

$$F(x,y,z,w)=x(5+z+x^2)+y(-10+w-x^3).$$

It is easy to see that the matrix H(F:x,y;z,w) has rank 2 in the set U_1xU_2. Then

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

$$|(x-x')(5+z)+(y-y')(-10+w)+x^3-x'^3-x^3y+x'^3y'|.$$

The Taylor series expansion of $x^3-x'^3-x^3y+x'^3y'$ around the point (x',y') is

$$3x'^2(1-y')(x-x')+(-x'^3)(y-y')+$$

$$1/2[6x_0(1-y_0)(x-x')^2+12x_0(x-x')(y-y')]$$

where (x_0,y_0) is a point on the line segment from (x,y) to (x',y'). Denote (x,y) by v and (x',y') by v'. The functions $3x'^2(1-y')$ and $(-x'^3)$ are the first derivatives of the cubic part of the expansion of F(x,y,z,w) and therefore these functions have limit zero as (x',y') approaches (0,0). The values $|6x_0(1-y_0)|$ and $|12x_0|$ are bounded above in the set U_1xU_2 by some number N. Furthermore

$$(x-x')^2\le P|v-v'|^2$$

and

$$|x-x'|\ |y-y'|\le P|v-v'|^2$$

for some positive real number P. Therefore

$$|1/2[6x_0(1-y_0)(x-x')^2+12x_0(x-x')(y-y')]|\leq$$
$$P'|v-v'|^2$$

for some positive real number P'. Then

$$|F(x,y,z,w)-F(x',y',z,w)|\geq$$
$$\left|\,\left|\frac{x-x'}{|v-v'|}(5+z+3x^2(1-y'))+\frac{y-y'}{|v-v'|}(-10+w-x^3)\right|-(P'|v-v'|)\,\right|$$

Choose a neighborhood U'_1 of (0,0) in U_1 so small that $|3x'^2(1-y')|$ and $|-x'^3|$ are both bounded by r/8 in that neighborhood and so that $P'|v-v'|<r/8$ in U'_1.

Then

$$\left|\ \left|\frac{(x-x')}{|v-v'|}(5+z+3x'^2(1-y'))+\frac{(y-y')}{|v-v'|}(-10+w-x'^3)\right|-\right.$$
$$\left.(P'|v-v'|)\ \right|\ |v-v'|\ \geq$$
$$\left|\ \left(\left|\ \left|\frac{(x-x')}{|v-v'|}(5+z)+\frac{(y-y')}{|v-v'|}(-10+w)\right|-\right.\right.\right.$$
$$\left.\left.r/4\right|\right)\left|-r/8\right|\quad |v-v'|$$

because

$$\left|r/8\frac{(x-x')}{|v-v'|}+r/8\frac{(y-y')}{|v-v'|}\right|=$$
$$|r/8\cos(\theta)+r/8\sin(\theta)|$$

for some $0\leq\theta\leq 2\pi$, and

$$|r/8\cos(\theta)+r/8\sin(\theta)|\leq r/4.$$

For each v-v', set

$$(\cos(\Theta(v-v')),\sin(\Theta(v-v')))=(v-v')/|v-v'|.$$

The inequality

$$|5\cos(\theta)-10\sin(\theta)|<3r/4$$

defines an open set I in R. Define a function

$$\chi(\theta)= \begin{cases} 1 \text{ if } \theta\in I \\ 0 \quad \text{otherwise} \end{cases}$$

Set

$$z(v-v')=(r)\chi(\theta(v-v'))\cos(\theta(v-v'))$$

and set

$$w(v-v')=(r)\chi(\theta(v-v'))\sin(\theta(v-v')).$$

Then

$$\left| \frac{(x-x')}{|v-v'|}(5+z)+\frac{(y-y')}{|v-v'|}(-10+w) \right|=$$

$$|\cos(\theta)(5+\chi(\theta)\cos(\theta))+\sin(\theta)(-10+\chi(\theta)\sin(\theta))|\geq$$

$$3r/4 \text{ if } \theta\notin I,$$

and

$$\left| \left|\left(\left|\frac{(x-x')}{|v-v'|}(5+z)+\frac{(y-y')}{|v-v'|}(-10+w)\right| \right|\geq\right.\right.$$

$$\left| \, |(r)\cos(\theta)^2+(r)\sin(\theta)^2)|- \right.$$

$$\left. |5\cos(\theta)-10\sin(\theta)| \, \right|\geq$$

$$| \; r-3r/4 \; |=r/4.$$

Then

$$|F(\ x,y,z,w\ )-F(\ x',y',z,w\ )|\geq$$

$$|(r/4)-(r/8)| \; |v-v'|.$$

We now give the general proof.

Proof. Set $Y=X_{<-i>}$ and denote X_i by X. Choose a point (p,q) in X x Y. In a neighborhood U x V of the point (p,q) suppose that X has coordinates $(x_1,...,x_m)$, $m=d(i)$, and that Y has coordinates $y_1,...,y_n$. That is, we can assume that U x V is mapped homeomorphically onto a neighborhood of the origin in $R^m \times R^n$ by the map, which we denote by (x,y), that carries (u,v) to $(x_1(u),....,y_n(v))$. The matrix H(F:z,w)[0,0] has rank m, because we have assumed that F is differentiably separable. Set $F^* = F\cdot(x,y)^{(-1)}$. It follows that F* is differentiably separable because the condition of differentiable separability is coordinate free. To lighten the notation, and at the risk of very little confusion, denote F* by F.

It follows that the coordinates in X and Y can be chosen compact neighborhoods U′ of p and V′ of q so that

$$\partial^2 F/\partial z_i\, \partial w_j\ (0,0) = \delta(i,j)$$

where $\delta(i,j)$ is Kronecker's delta function. We can now fix y and expand F(x,y) around the point (x′,y) using Taylor's Theorem (c.f. [4], p.200). Then

$$F(x,y) =$$

$$F(x',y) + \nabla F(x',y)\cdot(x-x') + \Theta(x^*,y)$$

where for some positive real number N,

$$|\Theta(x^*,y)| < N|x-x'|^2$$

in some compact neighborhood U" x V" of the point (0,0). The $\theta(x^*,y)$ is the remainder term of the Taylor series expansion, and x* is some point on the line segment from (x',y) to (x,y). The expression

$$\nabla F(x',y)\cdot(x-x')=$$

$$\Sigma_j\ \partial F/\partial x_j(x',y)(x_j-x'_j).$$

Expand the expression $\partial F/\partial x_j(x',y)$ around the point $(0,y)$. It follows that

$$\partial F/\partial x_j(x',y)=$$

$$\partial F/\partial x_j(0,y)+\Sigma_k\ \partial^2F/\partial x_j\partial x_k(0,y)x'_k+\Phi(x'',y)$$

where for a positive real number N' such that for all $1\le j\le m$,

$$|\Phi_j(x'',y)|<N'|x''|^2\le N'|x'|^2.$$

Expand $\partial F/\partial x_j(0,y)$ around the point $(0,0)$. It follows that

$$\partial F/\partial x_j(0,y)=\partial F/\partial x_j(0,0)+$$

$$\Sigma_k\ \partial^2F/\partial x_j\ \partial y_k(0,0)y_k+\Phi'_j(0,y^*),$$

where for some positive real number N'',

$$|\Phi'_j(0,y^*)|<N''|y|^2\le N''|y|^2.$$

Then

$$|F(x,y)-F(x',y)| \geq$$

$$\Big| \Big| \sum_j (\partial F/\partial x_j(0,0)+y_j+\sum_k \partial^2 F/\partial x_j \partial x_k(0,y)x_k$$

$$+\varphi_j(0,y*)+\varphi(x'',y))(x_j-x_j)\Big| - |\theta(x*,y)|\Big|.$$

The expressions $|\partial^2 F/\partial x_j \partial x_k(0,y)|$ are all bounded on the set V" by a real number T>0. Set

Λ=Max (N,N′,N",T).

Choose R>0 so small that the following conditions are are satisfied:

(i) the ball of radius R is contained in V",

(ii) $\Lambda(m^2+2)R^2<R/16$.

Choose r>0 so small that a ball of radius r around (0,0) lies in V" and such that $r<\min(1,R^2)$ and such that 2Nr<R/16.

Then

$$\Big| \sum_j (\partial F/\partial x_j(0,0)+y_j+\sum_k \partial^2 F/\partial x_j \partial x_k(0,y)x_k+\phi_j(0,y*)+$$

$$\phi(x'',y))(x_j-x_j)\Big| \geq$$

$$\Big|\Big|\sum_j [\partial F/\partial x_j(0,0)+y_j]\frac{x_j-x_j}{|x-x'|})\Big|(|x-x'|)-$$

$$\Big|\sum_k \partial^2 F/\partial x_j \partial x_k(0,y)x_k+\phi_j(0,y*)+\phi(x'',y))\Big|\Big|.$$

The vector $(x-x')/|x-x'|=v(x-x')$ is a unit vector. set

$(\partial F/\partial x_1(0,0),\ldots,\partial F/\partial x_m(0,0))=\Omega$.

Denote by S the unit sphere in R^m. The inequality $|\Omega \cdot s|<3R/4$ defines an open set I in the unit sphere S. Define a function $\chi(s)$, for $s \in S$ by the equation:

$$\chi(s)=\begin{cases} 1 & \text{if } s\in I \\ 0 & \text{otherwise.} \end{cases}$$

Set

$(y_1(v(x-x')),\ldots,y_m(v(x-x')))=$

$R\chi(v(x-x'))v(x-x')$.

If $v(x-x')\notin I$ and $|x|<r$, $|x'|<r$, then

$$\left|\,\left|\sum_j [\partial F/\partial x_j(0,0)+y_j]\frac{x_j-x_j}{|x-x'|}\right|\,(|x-x'|)-\right.$$

$$\left|\sum_k \partial^2 F/\partial x_j x_k(0,y)x_k+\varphi_j(0,y*)+\varphi(x'',y)\right|\,\Big|\geq$$

$$\left|\,\left|\sum_j (\partial F/\partial x_j\frac{x_j-x_j}{|x-x'|})\right||x-x'|\,\right|-\left|\sum_j (\sum_k \partial^2 F/\partial x_j\partial_k(0,y)x_k+\right.$$

$$\left.\varphi_j(0,y*)+\varphi(x'',y))\right|\,\left|\frac{x_j-x_j}{x-x'}\right|\,|x-\dot{x}'|\,\Big|\geq$$

$$|\,3R/4-m(m\Lambda r+\Lambda R^2+\Lambda r^2)\,|\,|x-x'|$$

But $r^2<R$ and $r<R^2$, therefore $m(m\Lambda r+\Lambda R^2+\Lambda r^2)<\Lambda(m^2+2)R^2\leq R/16$. It follows that

$$|F(x,y)-F(x',y)|\geq$$

$$|\,|3R/4-R/16||x-x'|-N|x-x'|^2\,|\geq$$

$$|5R/8-R/8||x-x'|,$$

because

$$|\Theta(x*,y)|<N|x-x'|^2<$$

$2Nr|x-x'|<(R/16)|x-x'|<$
$(R/8)|x-x'|$.

If $v(x-x')\in I$, if $|x|<r$, and if $|x'|<r$, then

$$\Big|\,\Big|\sum_j [\partial F/\partial x_j(0,0)+y_j]\Big(\frac{x_j-x_j)}{|x-x'|}\Big)\Big|\;|x-x'|-$$

$$\Big|\sum_k \partial^2 F/\partial x_j\partial x_k(0,y)x_k+\varphi_j(0,y*)+\varphi(x'',y))\Big|\,\Big|=$$

$$\Big|\,\Big|\Big\{\sum_j [\partial F/\partial x_j(0,0)+R\frac{x_j-x_j}{|x-x'|}]\frac{(x_j-x_j}{|x-x'|}\Big\}\Big|\;|x-x'|-$$

$$\Big|\sum_k \partial^2 F/\partial x_j\partial x_k(0,y)x_k+\varphi_j(0,y*)+\varphi(x'',y))\Big|\,\Big|\ge$$

$$\Big|\,\Big|\sum_j \frac{R(x_j-x_j)^2}{|x-x'|}\Big|\;|x-x'|-\Big|\sum_j (\partial F/\partial x_j(0,0)\frac{(x_j-x_j)}{|x-x'|}\Big|\;|x-x'|-$$

$$\Big|\sum_k \partial^2 F/\partial x_j\partial x_k(0,y)x_k+\varphi_j(0,y*)+\varphi(x'',y))\Big|\,\Big|\ge$$

$$\big|R|x-x'|-3R/4|x-x'|-m(m\Lambda r+\Lambda R^2+\Lambda r^2)|x-x'|-(r/16)|x-x'|\big|\ge$$

$$|R-3R/4-R/16-R/16|\;|x-x'|=(r/8)|x-x'|.$$

Therefore, we set $M=R/8$.▒

We use the separation properties that have been established in Theorem 10.1 to estimate the limiting value of the computing time required for ϵ-approximations of a differentiably separated function when there is a precise relation between δ and ϵ. The next lemma (and definition) states the relation.

Lemma 10.2. Suppose that $g(\epsilon)$ is a continuously differentiable function of a real variable ϵ, defined on an interval around 0. Assume that g satisfies the following conditions:

(i) for each $\epsilon>0$, $0<g(\epsilon)<\epsilon$,

(ii) $\frac{dg}{d\epsilon}(0)\neq 0$.

If K, M, and N are positive real numbers, and if

$n(\epsilon)=\text{int}[N/g(\epsilon)]$, then

$$\lim_{\epsilon\to 0} \log_{n(\epsilon)}\left[\frac{\text{int}[K/g(\epsilon)]}{\text{INT}[\epsilon/Mg(\epsilon)]}\right]=1.$$

A function $g(\epsilon)$ that satisfies the conditions of this lemma will be called a delta-epsilon function.

Proof. Set $L(\epsilon)=\log_{n(\epsilon)}\left[\frac{\text{int}[K/g(\epsilon)]}{\text{INT}[\epsilon/Mg(\epsilon)]}\right]$.

The definitions of int and INT imply that

$$\left[\frac{N}{g(\epsilon)}\right]\leq \text{INT}\left[\frac{N}{g(\epsilon)}\right]\leq\left[\frac{N}{g(\epsilon)}\right] +1 \text{ and}$$

$$\left[\frac{K}{g(\epsilon)}\right]- 1 \leq \text{INT}\left[\frac{N}{g(\epsilon)}\right] \leq \left[\frac{K}{g(\epsilon)}\right].$$

with similar inequalities for

$$\text{INT}\left[\frac{\epsilon}{Mg(\epsilon)}\right].$$

Because $\log_a b=(\ln a)/(\ln b)$, it follows that if we set

$$I(\epsilon) = \frac{\ln\left[\dfrac{(K/g(\epsilon)) - 1}{(\epsilon/Mg(\epsilon)) + 1}\right]}{\ln\left[(N/g(\epsilon)) + 1\right]}$$

and set

$$II(\epsilon) = \frac{\ln\left|\dfrac{(K/g(\epsilon))}{(\epsilon/Mg(\epsilon))}\right|}{\ln\left[N/g(\epsilon)\right]}$$

then $I(\epsilon) \le L(\epsilon) \le II(\epsilon)$. However,

$$\lim_{\epsilon \to 0} I(\epsilon) = \lim_{\epsilon \to 0}\left[\frac{\ln\left[\dfrac{(K-g(\epsilon))\ M}{\epsilon + Mg(\epsilon)}\right]}{\ln\left[\dfrac{N+g(\epsilon)}{g(\epsilon)}\right]}\right].$$

Because $\epsilon + Mg(\epsilon)$ and $g(\epsilon)$ both have limit zero as e approaches 0, we can apply L'Hospital's Rule to compute the limit. Therefore $\lim_{\epsilon \to 0} I(\epsilon) =$

$$\lim_{\epsilon\to 0}\left[\frac{-\epsilon g'(\epsilon)-K-KMg'(\epsilon)+g(\epsilon)}{(K-g(\epsilon))\left[1+M\frac{g(\epsilon)}{\epsilon}\right]}\right]\left[\frac{(N+g(\epsilon))\frac{g(\epsilon)}{\epsilon}}{-g'(\epsilon)N}\right]=$$

$$\left[\frac{-K-KM\,g'(0)}{K(1+Mg'(0))}\right]\left[\frac{Ng'(0)}{-g'(0)N}\right]=1,$$

because

$$\lim_{\epsilon\to 0} g(\epsilon)/\epsilon = g'(0).$$

Similarly

$$\lim_{\epsilon\to 0} II(\epsilon) = 1.$$ ▒

Lemma 10.3. Suppose that $F: X_1 x\ldots x X_n \to R$ is a continuously differentiable function on $U_1 x\ldots x U_n$ where each U_i is a nonempty open subset of X_i. Suppose that for each i, $1\le i\le n$, K_i is a nonempty compact subset of U_i. There exists a delta-epsilon function $g(\epsilon)$ such that for each ϵ sufficiently small, and for each $x, x'\in K_1 x\ldots x K_n$, if $|x-x'|<g(\epsilon)$, then

$$|F(x)-F(x')|<\epsilon.$$

Proof. Denote by S the unit sphere in $X_1 x\ldots x X_n$. For each $x \in U_1 x\ldots x U_n$ and each $v \in S$, set $D(x,v)=\left|\frac{\partial F}{\partial t}(x)\right|$. The function $D(x,v)$ is continuous from $K_1 x\ldots x K_n x\ S$ to R. It follows that there is a real number $B>0$ such that $D(x,v)<B$ for all $x\in K_1 x\ldots x K_n$ and $v\in S$. Clearly we can choose $B>1$. Suppose that $y, y'\in K_1 x\ldots x K_n$ such that $y\ne y'$.

Set

$$v^* = \frac{y-y'}{|x-x'|} \qquad ,$$

set

$$q(t) = y' + tv^*,$$

Let

$$f_L(t) = F(q(t)).$$

Then,

$$|F(y) - F(y')| = \left|\frac{d}{dt}F_L(a)\right| . |y-y'|$$

for some a, $0 < a < |y-y'|$. But

$\left|\frac{d}{dt}F_L\right|$ is bounded by B, therefore

$$|F(y) - F(y')| < B . |y-y'| .$$

We can choose as the delta-epsilon function,

$g(\epsilon) = \epsilon/B$. ▒

<u>Definition 10</u>. Suppose that X_j, $1 \le j \le n$ are differentiable manifolds of dimensions $d(1), \ldots, d(n)$, respectively. Suppose that $F: \prod_1^n X_j \longrightarrow R$ is a differentiable function. Assume that $(p_1, \ldots, p_n) \in \prod_1^n X_j$, suppose that for each $1 \le j \le n$, U_j is a coordinate neighborhood of p_j, and suppose that $\varphi_j = (\varphi_{j\ 1}, \ldots, \varphi_{j\ d(j)}): U_j \longrightarrow R_{d(j)}$ is a set of local coordinates at p_j. If for each $1 \le j \le n$, L_j is a regular lattice on $R^{d(j)}$, then the set $(\prod \varphi_j)^{-1}(\prod L_j)$ will be called a <u>regular lattice</u> on $\prod X_j$ along the coordinates $\varphi_{1\ 1}, \ldots, \varphi_{n\ d(n)}$ in the

coordinate neighborhood $\prod U_j$. The mesh of the lattice $(\prod\varphi_j)^{-1}(\prod L_j)$ is the <u>mesh of the lattice</u> $\prod L_j$.

<u>Definition 10.4</u>. Suppose that for $1 \le j \le n$, X_j is a differentiable manifold, and suppose that $\varphi_i : U_i \text{--->} R^{d(i)}$ is a local coordinate system for X_i in the neighborhood of a point p_i. If $(\prod\varphi_i)^{-1}(\prod L_i) = L$ is a regular lattice on $\prod X_i$ along the coordinates $\varphi_{1\ 1}, \ldots, \varphi_{n\ d(n)}$, then a function $f_L : \prod U_i \text{--->} R$ is an <u>ϵ-approximation of F in the neighborhood $\prod U_i$</u>, if the function $f_L \cdot (\prod\varphi_i)^{-1}$ is an ϵ-approximation of the function $F \cdot (\prod\varphi_i)^{-1}$ in the set $\prod\varphi_j(U_j)$.

<u>Theorem 10.2</u>. Suppose that $F : X_1 x \ldots x X_n \text{-->} R$ is a continuously differentiable function from the product of Euclidean spaces X_i to the real numbers. Suppose that for each i, the space X_i has standard basis $\{e_{(i\ j)}\}$, $1 \le j \le d(i)$.

Assume:

(i) for each positive real number ϵ, sufficiently small, $L_i(\epsilon)$ is a regular lattice in X_i along the basis $\{e_{(i\ j)}\}$;

(ii) there is a delta-epsilon function $g(\ \epsilon\ )$ (for ϵ sufficiently small) and for each i, a compact set K_i with nonempty interior contained in X_i such that for each $x, x' \in K_1 x \ldots x K_n$, if $|x-x'|<g(\epsilon)$, then

$$|F(x)-F(x')|<\epsilon;$$

(iii) the lattice $L_i(\epsilon)$ has mesh $g(\epsilon)$ for each i and each ϵ sufficiently small;

(iv) for each $1 \le i \le n$ there is an open set U_i in X_i that contains K_i and there is a (nonempty) submanifold S_i of U_i such that F has separator functions W_i along the sets $\{S_i\}$ in the neighborhood $U_1 x \ldots x U_n$;

(v) there is a lattice function $f^{(\epsilon)}: \prod_{i=1}^{n} L_i(\epsilon) \dashrightarrow L(\epsilon)$ that is an ϵ-approximation of F;

(vi) there is a real number $K>0$, such that for $d(\epsilon)=\mathrm{int}[K/g(\epsilon)]$, there is an $(r,d(\epsilon))$-network $C(\epsilon)$ that computes the function $f^{(\epsilon)}$.

If $C^{(\epsilon)}$ computes $f^{(\epsilon)}$ in time $T(\epsilon)$, then

$$\lim_{\epsilon \to 0} T(\epsilon) \ge \mathrm{INT}[\log_r(\Sigma \dim S_i)].$$

Proof. Fix an integer i, $1 \le i \le n$. Choose a point $s \in \prod(U_i)$. Suppose that $\dim S_i = \sigma_i$. Because the S_i are submanifolds of X_i we can choose a coordinate neighborhood of s so small, and if necessary a re-indexing of the basis elements $\{e_{i\ k}\}$, such that for each j there is a coordinate system $\{x_{j\ 1},\ldots,x_{j\ d(j)}\}$ at s with coordinate lines in the direction $\{e_{j\ 1},\ldots,e_{j\ \sigma(j)}\}$ and such that the projection of S_j into the linear subspace P_j with equations

$$x_{j\ \sigma(j)\ +1}=\ldots=x_{j\ d(j)}=0$$

is a diffeomorphism; that is, the projection is a diffeomorphism in the neighborhood U_j. By Theorem 10.1 F has separator functions W_i in the open set $U_1x\ldots xU_n$ along the subsets $\{S_i\}$. Therefore by Definition 10.1 there is a real number M>0, such that for each $y\neq y'$ in $\prod(U_i\cap S_i)$,

$$|F(\ y,W_i(\ y,y'\ )\ )-F(\ y',W_i(\ y,y'\ )\ )|\geq M|y-y'|.$$

For a sufficiently small S, we can choose for each $1\leq i\leq n$ a cube B_i of side length S that is contained in K_i and has sides parallel to the basis elements $e_{(i\ j)}$ and such that the vertices of the cube are vertices of the lattice $L_i(\epsilon)$.

Suppose that y and y′ are vertices of $L_i(\epsilon)\cap P_i$ that $|y-y'|>4\epsilon/M$. Because the projection of S_i to P_i is a diffeomorphism there are points q(y) and q(y′) that lie on S_i that project onto y and y′, respectively. The point q(y) lies in a rectangle whose principal vertex we denote by v. Denote by v′ the principal vertex of the cube that contains q(y′). Suppose the point $W_i(\ y,y'\ )$ lies in a cube of $X_{<-i>}$ that has principal vertex w. Then the point $y\int_i w$ is a lattice point of $\prod L_i(\epsilon)$. By assumption, f^ϵ is an ϵ-approximation of F. It follows that

$$|f^\epsilon(y\int_i w)-F(y\int_i w)|<\epsilon$$

because $(y\int_i w)$ lies in the cube with principal vertex $y\int_i w$. Similarly,

$$|f^\epsilon(y'\int_i w)-F(y'\int_i w)|<\epsilon.$$

Therefore,

$$|f^\epsilon(y\int_i w)-f^\epsilon(y'\int_i w)|=$$
$$|f^\epsilon(y\int_i w)-F(y\int_i w)+F(y\int_i w)-$$
$$f^\epsilon(y'\int_i w)-F(y'\int_i w)+F(y'\int_i w)| >$$
$$| \ |F(y\int_i w)-F(y'\int_i w)|-$$
$$|f^\epsilon(y\int_i w)-F(y\int_i w) -f^\epsilon(y'\int_i w)+$$
$$F(y'\int_i w)| \ | \geq$$
$$|4\epsilon-2\epsilon|=2\epsilon.$$

The lattice $L_i(\epsilon)$ has mesh $g(\epsilon)$, and therefore the number of vertices along one side of the cube B_i in an interval of length S is int[$S/g(\epsilon)$]. Along any one axis of the lattice $L_i(\epsilon)$, the distance between the j^{th} vertex and the $j+h^{th}$ vertex is $hg(\epsilon)$. If $hg(\epsilon)>4\epsilon/M$, vertices are $4\epsilon/M$ units apart. Set $s=INT[4\epsilon/(Mg(\epsilon))]$. Along each side of the cube B_i choose every s^{th} vertex. Along each such side, the number of vertices chosen is

$$D=\text{int}\left[\frac{\text{int}\left[\frac{S}{g(\epsilon)}\right]}{\text{INT}\left[\frac{4\epsilon}{Mg(\epsilon)}\right]}\right].$$

The number of vertices we have chosen in the cube B_i is $D^{d(i)}$, and all of these vertices are at least $(4\epsilon/M)$ units apart. Set $n(\epsilon)=\text{int}[K/g(\epsilon)]$. The minimum computing time to compute f^{ϵ} using an $(r,\text{int}[S/g(\epsilon)])$ network is then $\text{INT}[\log_r\{\Sigma \log_{n(\epsilon)} D^{d(i)}\}]$. To complete the proof of the assertion, it will suffice to show that

$$\lim_{\epsilon \to 0} \log_{n(\epsilon)}\left[\frac{\text{int}\left[\frac{S}{g(\epsilon)}\right]}{\text{INT}\left[\frac{4\epsilon}{Mg(\epsilon)}\right]}\right]=1.$$

But this is the conclusion of Lemma 10.2.▒

Chapter XI

A Limit Theorem for C^n-Networks

In this chapter we analyze the time needed to compute a C^n function F by C^n (2,1)-networks as a limit of times taken by finite networks that compute finite approximations to F. In the limiting process studied here the structure of the approximating networks remains fixed while the size of the alphabet is allowed to vary. This may be interpreted to say that the same algorithm is used to compute the finite approximating functions and the limiting function, while increasingly many symbols are used to encode the finer approximations as we pass to the limit, much as the number of positions in rational approximations to real numbers increases as we consider progressively finer measurements.

Theorem 11.1 states the result of interest, a result that helps to justify the use of continuous (or C^n) networks to represent computing. Two lemmas, Lemmas 11.1 and 11.2 used in the proof of the Theorem 11.1, are of a purely technical nature and are stated and proved following the proof of the theorem. The hypotheses of Lemma 11.1, which are part of the hypothesis of Theorem 11.1, can be satisfied, for example, by using polynomials for the approximating functions.

Theorem 11.1: Let $F:V_C$--->R, $V_C=R_C \times \ldots \times R_C$, where R_C is a compact neighborhood of zero in R and V_C is a compact neighborhood of zero in the Euclidean space V, be a function satisfying F(0) = 0 that is computed by a continuous (resp. C^n) (2,1)-network in time t. Suppose further that if a continuous (resp. C^n) (2,1)-network computes F in time t',then $t' \geq t$.

For j = 1, 2, ..., let $\epsilon_j > 0$ be such that $\epsilon_j > \epsilon_{j+1}$ and ϵ_{j+1} and $\epsilon_j \to 0$ as $j \to \infty$. Let $\{C^j\}$ be a sequence of finite loop free $(2,d_j)$-networks such that C^j computes an ϵ_j-approximation to F in a bounded neighborhood of 0, in time t_j, and suppose that the modules of C^j can be approximated at lattice points by continuous (resp. C^n) functions, as described in the hypotheses of Lemma 11.1[10]. Then

$$\tau = \lim\inf \{t_j\}$$

satisfies the inequality $\tau \geq t$.

Proof: Let $\{\epsilon_j\}_{j=1}^{\infty}$ be a sequence of real numbers where $\epsilon_j > 0$, $\epsilon_j > \epsilon_{j+1}$ and ϵ_j--->0 as j--->∞. Let $F_j:L_j \times \ldots \times L_j$--->$L_j$ be an $\epsilon_j/2$-approximation of F, where L_j is the lattice of a rectangular decomposition of R_C.

---

10) By Corollary C.1 we can, without loss of generality, confine attention to loop free networks.

For each j, and ϵ_j, let C_j be a $(2,d_j)$-network with alphabet L_j that computes F_j in time t_j. Since $\tau_j>0$, $\tau = \liminf \tau_j \geq 0$.

If $\tau \geq t$, there is nothing to prove. So, suppose $\tau < t$. Because the τ_j and hence τ are integers, there is an infinite subsequence $\{C_{j_q}\} \subseteq \{Cj\}$ of networks that compute F_{j_q} in time τ.

For given τ, the number of binary trees of depth τ is finite. Therefore, there must be an infinite subsequence $\{C_{j_{q_r}}\} \subseteq \{C_{j_q}\}$ of $(2,d_j)$-networks each of

which computes $F_{j_{q_r}}$ and all of which have the same

graph. To simplify notation, let us call this subsequence $\{C_j\}$ and correspondingly the sequence of approximations $\{F_j\}$. Thus, $\{C_j\}$ is an infinite sequence of $(2,d_j)$-networks, each with the same graph, that compute F_j in time τ.

Since all the networks C_j have the same graph, we can identify unambiguously modules in the same position in different networks. Let $G^i{}_j$ be the module (function) in position i in network j.

Given the functions $G^i{}_j$: $L_j \times L_j \text{---> } L_j$,

$i=1,\ldots,q$, $\epsilon_j>0$, under the hypothesis of Lemma 11.1 there exist continuous (resp. C^n) functions

$P^i{}_j: R_C \times R_C \text{--->} R_C$ such that

$$| P^i{}_{n(\epsilon_j)}(\ell) - G^i{}_j(\ell)| < \epsilon_j/4 ,$$

for $\ell \in L_j \times L_j$.

Under the hypothesis of Lemma 11.1 for each $i=1,\ldots,q$, the sequence of continuous (resp. C^n) functions $\{P^i{}_{n(\epsilon_j)}\}_{j=1}^{\infty}$ has a uniform limit that is continuous (resp. C^n). Thus the sequence of networks C_j converges to a network C' whose graph is the same as the common graph of the C_j and whose modules are the limits of the functions $P^i{}_j$ as $j\text{--->}\infty$. Thus, the network C' is a continuous (resp. C^n) (2,1)-network. Furthermore, the network C' computes F. To see this, let C'_j be the network that results from substituting the function $P^i{}_j$ in place of the function $G^i{}_j$ in C_j. Denote by

$F'_j:V_C\text{--->}R$ the function computed by C'_j.

Since for each ϵ_j, we may choose the functions $P^i{}_j$ such that

$$| P^i{}_j(\ell) - G^i{}_j(\ell)| < \epsilon_j/4,$$

it follows from Lemma 11.2 that for k sufficiently large, $|F'_k - F_k'| < \epsilon_j/2$.

Then,

$$| F'_k - F| \le |F'_k - F_k| + | F_k - F | < \epsilon_j/2 + \epsilon_j/2 = \epsilon_j.$$

Finally, the limiting network C′ computes F in time τ, because its graph is that of a loop free network whose delay is τ.

Thus, the limiting network C′ is a continuous (resp. C^n) (2,1)-network that computes F in time $\tau<t$. But this is impossible, because by hypothesis, t is the minimum delay among all such (2,1)-networks that compute F. This concludes the proof. ▒

The limiting argument in the proof of Theorem 11.1 uses a sequence of continuous (C^n) functions that approximate the modules of the finite networks at lattice points. Lemma 11.1 gives conditions under which the values of such an approximating function cannot differ much from the finite function everywhere on the rectangles of the lattice decomposition.

<u>Lemma 11.1</u>:

Let $\{\epsilon_j\}$ be a sequence of positive numbers decreasing to zero as j tends to infinity. For each ϵ_j let L_j be a lattice decomposition of R_C (a compact neighborhood of zero in R) such that

a) $| P_j(\ell) - G_j(\ell)| < \epsilon_j/4$, for $\ell \in L_j \times L_j$,

where the functions

P_j: $R_C \times R_C$--->R_C

form a sequence of continuous (resp. C^n) functions that converge equi-continuously to a continuous (resp. C^n)

function

$$P: R_C \times R_C \longrightarrow R_C$$

and

$$G_j: L_j \times L_j \longrightarrow L_j$$

is a (finite) function for each $j = 1, 2, \ldots,$ and

b) the mesh of the lattice L_j (c.f. Definition 9.7) decreases to zero as j tends to infinity. Then, for j sufficiently large, G_j is an $\epsilon_j/2$-approximation to P_j, i.e.

$$| P_j(x) - G_j [\ell(x)] | < \epsilon_j/2,$$

for all $x \in R_C \times R_C$.

<u>Proof of Lemma 11.1</u>

To show that for j sufficiently large, G_j is an $\epsilon_j/2$-approximation of P_j on $R_C \times R_C$ it suffices to show that if $x \in D_j[\ell(x)]$, then

$$| P_j(x) - G_j[\ell(x)] | < \epsilon_j/2.$$

Now,

$$|P_j(x) - G_j[\ell(x)] \leq | P_j(x) - P_j [\ell(x)] | + | P_j[\ell(x)] - G_j [\ell(x)] |.$$

By hypothesis

$$| P_j(x) - G_j [\ell(x)] | < \epsilon_j/4, \text{ for all } j.$$

Therefore it remains to show that the other term is small. Since P_j is uniformly continuous in $R_C \times R_C$, for every $\eta_j > 0$, there exists $\gamma_j(\eta_j) > 0$ such that

$$| x - \ell(x) | < \gamma_j(\eta j)$$

implies

$$| P_j(x) - P_j[\ell(x)] | < \eta_j.$$

Under assumption a) of the Lemma, P_j converges equicontinuously to a continuous (resp. C^n) function $P: R_c \times R_c \text{--->} R_c$. Since P is also uniformly continuous on $R_c \times R_c$, there is a function $\gamma: R \text{--->} R$ such that for every $\eta>0$,

$$|x - y| < \gamma(\eta) \text{ implies}$$
$$| P(x) - P(y) | < \eta.$$

Now,

$$| P_j(x) - P_j[\ell(x)] | =$$
$$| P(x) - P[\ell(x)] - P(x) + P_j(x) +$$
$$P[\ell(x)] - P_j[\ell(x)] | \leq$$
$$| P(x) - P[\ell(x)] | + | P(x) - P_j(x) | +$$
$$| P[\ell(x)] - P_j[\ell(x)] |.$$

Given η, there exists an integer $J(\eta/3)$ such that $j > J(\eta/3)$ implies

$$| P(x) - P_j(x) | < \eta/3$$

and

$$| P[\ell(x)] - P_j[\ell(x)] | < \eta/3.$$

Further

$$| x - \ell(x) | < \gamma(\eta/3)$$

implies

$$| P(x) - P[\ell(x)] | < \eta/3.$$

Now, let $\eta_j = \epsilon_j/4$ and let $J_j = J(\eta_j/3) + J(\eta_j/12)$. Further, define $\sigma: N \text{--->} N$ by the condition that $\sigma(j)$ is the smallest integer, k, such that $\gamma(\eta/3) > \delta_k$, where $\delta_k = |D_k|$, is the mesh of lattice L_k. Thus, corresponding to the sequences j = 1, 2 ... and $\{\epsilon_j\}$, there are the sequences $\{J_j\}$ and $\{\sigma(j)\}$ such that for all

$k > \max \{J_j, \sigma(j)\} = K(j)$

$$| P_k(x)-P_k [\ell(x)] |< \epsilon_j/4.$$

Hence for each j and all $k > K(j)$,

$$| P_k(x)-G_k[\ell(x)] | \leq$$
$$| P_k(x)-P_k[\ell(x)] | +$$
$$| P_k[\ell(x)]-G_k [\ell(x)] | \leq$$
$$\epsilon_j/4 + \epsilon_j/4= \epsilon_j/2.$$

This completes the proof. ▒

The sequence of finite networks generated by the construction in the proof of Theorem 11.1 computes a sequence of finite functions. The sequence of continuous (C^n) networks generated by approximating the finite modules of the first sequence by continuous (C^n) modules also computes a sequence of functions. Lemma 11.2 establishes that these two sequences of functions converge to a common limit.

Lemma 11.2

Let $\{\epsilon_j\}$ be a sequence of positive numbers decreasing to zero as j tends to infinity. For each ϵ_j, let L_j be a lattice decomposition of R_c and let C_j

be a finite $(2,d_j)$-network, with alphabet L_j, modules

$$G^i_j: L_j \times L_j \text{--->} L_j \ , \quad i=1,\ldots,q$$

and a common loop free digraph for all j, which computes a (finite) function

$$F_j:L_j \times\ldots\times L_j \text{--->} L_j.$$ [11]

Let C'_j be a (2,1)-network with the same digraph as C_j, whose module in position i is[12]

$$P^i_j: R_C \times R_C \text{--->} R_C$$

in place of G^i_j where for i = 1,...,q, P^i_j and G^i_j satisfy the hypotheses of Lemma 11.1, and where C'_j computes a function

$$F'_j: R_C \times\ldots\times R_C \text{--->} R_C.$$

---

[11] Let L be the lattice of a rectangular decomposition of R_C, a compact neighborhood of zero in R,

and let F:L x L--->L be a finite function.

There is a (2,d)-network C that computes F, where d is equal to the number of points in L. In that case the alphabet used by C can be identifies with the lattice L and the modules of C with functions from L x L to L.

If C uses an alphabet A such that g:L--->A is a one-to-one encoding of L onto A (both sets having d elements), and if G~ are the modules of C, then the functions

$$G(\ v\ )(\ell_1,\ell_2)=G\sim(g(\ell_1),g(\ell_2))$$

are the modules of the corresponding (2,d)-network with L as alphabet. It is straightforward to show that if

α:A x A--->A is everywhere computed by C and

F:L x L--->L is computed by the network with modules G(v) and the same digraph as C, then F and α satisfy the relation

$$g(\ F(\ \ell_1,\ell_2\ )\ )=\alpha(\ g(\ \ell_1\ ),g(\ \ell_2\ )\ ).$$

[12] The module in position i in the network C_j is well-defined because all networks are finite and have the same digraph.

For each ϵ_j there exists an integer $K(j)=K(\epsilon_j)$, such that for all $k > K(j)$ and for all $x \in R_C \times \ldots \times R_C$

$$| \; F'_k(x)-F_k[\ell(x)] \; |< \epsilon_k/2.$$

<u>Proof of Lemma 11.2</u>

We shall give the argument for the case of a loop-free network of delay 2. The same argument applies in general; the notation is less complicated and the argument more easily followed in the delay 2 case. In that case the domain of F_j is (at least) four dimensional. Then, let

$$x = (x_1,x_2,x_3,x_4)$$

and $\ell(x) = (\ell(x_1),\ell(x_2),\ell(x_3),\ell(x_4))$.

We suppose the network for F_j to be

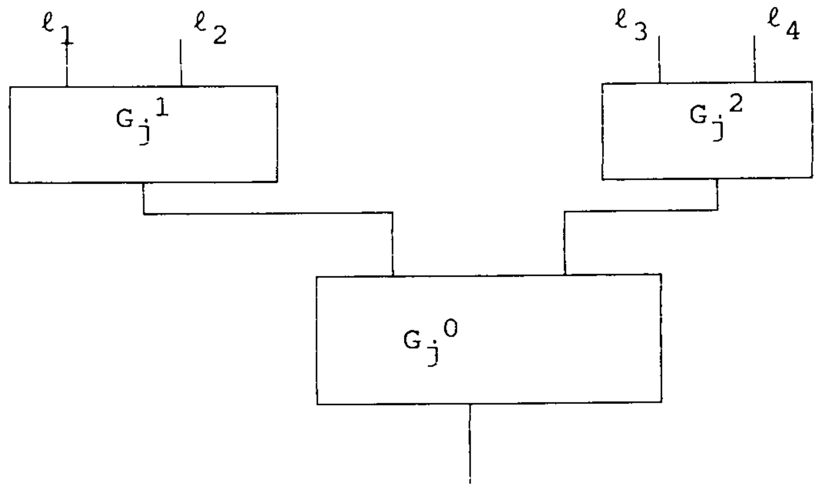

Figure 11.1

and similarly, for F'_j to be

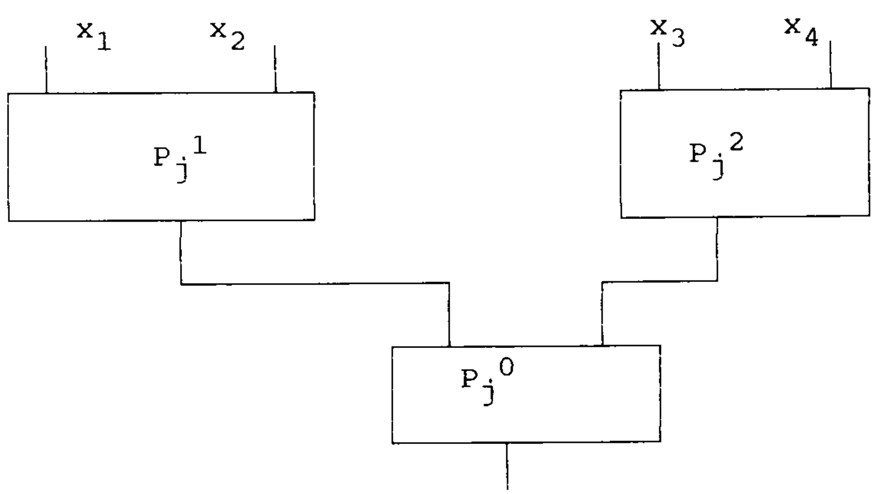

Figure 11.2

We shall write $\ell^1(x)=[\ell^1(x),\ell^2(x)]$ and $\ell^2(x)=[\ell^3(x),\ell^4(x)]$ and $x^1=(x_1,x_2)$, $x^2= (x_3,x_4)$. Then:

$$F'_j(x)=P^0_j [P^1_j(x_1),P^2_j(x_2)]$$

and

$$F_j[\ \ell(x)\ ]=$$
$$G^0{}_j[\ G^1{}_j[\ \ell^1(\ x\ )\ ],G^2{}_j[\ \ell^2(\ x\ )\ ]\ ].$$

Thus,

$$|\ F'{}_j(\ x\ )-F_j[\ \ell(\ x\ )\ ]\ |\ =$$
$$|\ P^0{}_j[\ P^1{}_j(\ x_1\ ),P^2{}_j(\ x_2\ )\ ]\ -$$
$$G^0{}_j[\ G^1{}_j[\ \ell^1(\ x\ )\ ],G^2{}_j[\ \ell^2(\ x\ )\ ]\ ]\ =$$
$$|\ P^0{}_j(\ u_j,v_j\ )\ -\ G^0{}_j(\ w_j,z_j\ )\ |\leq$$
$$|\ P^0{}_j(\ u_j,v_j\ )\ -\ P^0{}_j(\ w_j,z_j\ )\ \ |\ +$$
$$|\ P^0{}_j(\ w_j,z_j\ )\ -\ G^0{}_j(\ w_j,z_j\ )\ |,$$

where

$$(u_j,v_j)=[P^1{}_j(\ x_1\ ),P^2{}_j(\ x_2\ )],$$

and

$$(w_j,z_j)=[G^1{}_j(\ \ell^1(\ x\ )\ ),G^2{}_j(\ \ell^2(\ x\ )\ )].$$

Since $G^i{}_j$: L_j x L_j---> L_j, it follows that

$$\ell(w_j,z_j)\ =\ (w_j,z_j)\ \in L_j\ x\ L_j.$$

Hence $P^0{}_j$ and $G^0{}_j$ are defined at (w_j,z_j). By hypothesis,

$$|\ \ P^0{}_j(\ w_j,z_j\ )\ -\ G^0{}_j(\ w_j,z_j\ )\ |\ <$$
$$\epsilon_j/4\ <\ \epsilon_j/2$$

for all j. It remains to show that

$$|\ \ P^0{}_j(\ u_j,v_j\ )\ -\ P^0{}_j(\ w_j,z_j\ )\ \ |\ <\ \epsilon_j/4.$$

Since $P^0{}_j$ is uniformly continuous on R_C x R_C, for every $\eta_j>0$ there exists $\delta_j(\eta_j)>0$ such that

$$|\ (u_j,v_j)-(w_j,z_j)\ |<\ \delta_j(\eta_j)$$

implies

$$| P^0_j(u_j,v_j) - P^0_j(w_j,z_j) | < \eta_j.$$

Hence, it suffices to show that

$$| (u_j,v_j) - (w_j,z_j) | < \delta_j(\epsilon_j/4).$$

Now,

$$| (u_j,v_j) - (w_j,z_j) |^2 =(u_j-w_j)^2+(v_j-z_j)^2=$$
$$[P^1_j(x) - G^1_j(\ell(x))]^2 +$$
$$[P^2_j(x)-G^2_j(\ell(x))]^2.$$

It follows from Lemma 11.1 that for k sufficiently large, G^i_k is an $\epsilon_j/4$-approximation of P^i_j on $R_c \times R_c$ for each i, by taking $[\delta_j(\epsilon_j/4))/\surd 2]$ in place of $(\epsilon_j/4)$ in that proof, for all $k \geq K(j)$.[13)]

---

13) Since q is finite, a standard argument shows that there is a value K(j) that works for all i=1,...,q.

$$| P^i_k(x)-G^i_k(\ell(x))| \leq \delta_j(\epsilon_j/4)/\surd 2$$ for i=1,2.
Hence

$$| P^1_k(x)-G^1_k[\ell(x)] |^2 +|P^2_k (x^2)-G^2_k(\ell^2(x) |^2 \leq$$

$$\delta_j(\eta_j)^2/2 + \delta_j(\eta_j)^2/2 =[\delta_j(\eta_j)]^2.$$

It follows that

$$|(u_k,v_k)-(w_k,z_k)|=$$
$$\{ P^1_k(x^1)-G^1_k(\ell^1(x))^2 +(P^2_k(\ell^2(x))^2\}^{1/2} \leq \delta_j(\epsilon_j/4)\}.$$

Thus, it follows that for each j and hence each ϵ_j, there exists an integer $K(j)=K(\epsilon_j)$ such that $k>K(j)$ implies

$$| F'_k(x)-F_k[\ell(x)] | \leq \epsilon_j/4$$

uniformly in x.

Appendix A

Privacy preserving correspondences

Section A1. Privacy Preserving Correspondences

The assumption that a correspondence is privacy preserving is a very strong condition, independent of any continuity assumptions. It is the set theory of privacy preserving correspondences that we discuss in this appendix. One can find in the paper [11] a discussion of message spaces and mechanisms that realize differentiable functions. When mappings and correspondences are not required to satisfy topological conditions, some of the discussion becomes more transparent. We begin by analyzing a simple example.

Suppose that a function F defined on the product of two sets X_1 and X_2, where each X_i consists of three elements, takes the values 0 and 1. Label the points of the set X_1 as a,b,and c and label the points of the second set as e,f,and g. One can describe the function F easily using a matrix M=M(F) of 0's and 1's with rows indexed by a,b, and c and with columns labelled e,f, and g. The $(x,y)^{th}$ entry in the matrix M is the value of the function at the point $(x,y)\in X_1 x X_2$. For example, the matrix M(F)=

	e	f	g
a	1	1	0
b	1	1	1
c	0	1	1

represents a function F defined on $X_1 x X_2$, where F(a,f)=1, F(c,e)=0, etc. We then ask what correspondence μ from $X_1 x X_2$ onto a set M is a privacy preserving correspondence that can be used to realize F. The definition of privacy preserving correspondence (c.f. Definition A1.1) states that μ is privacy preserving if there are correspondences $\mu^i : X_i \text{--->} M$ such that for each $(x,y) \in X_1 x X_2$, $\mu(x,y) = \mu^1(x) \cap \mu^2(y)$. Furthermore, if μ realizes F (c.f. Definition A4.1) if there is a function h:M-->{0,1} such that for each $(x,y) \in X_1 x X_2$, h is constant on $\mu(x,y)$ and $h(\mu(x,y)) = F(x,y)$. We now attempt to realize the function F, given by the matrix M(F), with a privacy preserving correspondence $\mu : X_1 x X_2 \text{--->} M$. A reasonable candidate for the correspondence μ is the function F, itself, setting M={0,1}. One the function h:M--->{0,1} is the identity function. We now face the problem of deciding if the correspondence, or in this case the function μ=F, is privacy preserving. This might seem to lead to the disagreeable task of enumerating the possible correspondences

$\mu_i : X_i \dashrightarrow \{0,1\}$.

However, the requirement that

$$\mu(x,y)=\mu^1(x)\cap\mu^2(y) \text{ for all } (x,y)\in X_1 x X_2$$

imposes a special structure on the inverse correspondence $\mu^{-1}: M \dashrightarrow X_1 x X_2$ that is easy to check.

If m is a point in the set M, the set $\mu_1^{-1}(m)$ is a subset of X_1 and the set $\mu_2^{-1}(m)$ is a subset of X_2. Suppose $u\in\mu_1^{-1}(m)$ while $v\in\mu_2^{-1}(m)$. Then $m\in\mu_1(u)$ and $m\in\mu_2(v)$, so

$$m\in\mu_1(u)\cap\mu_2(v)=\mu(u,v).$$

That is, the set $\mu(u,v)$ contains the product

$$R=\mu_1^{-1}(u) x\ \mu_2^{-1}(v).$$

It is equally easy to see that R actually equals $\mu^{-1}(m)$. Indeed, if

$$(u',v')\in\mu^{-1}(m),$$

then

$$m\in\mu(u',v')=\mu_1^{-1}(u')\cap\mu_2^{-1}(v')$$

and therefore

$$u'\in\mu_1^{-1}(u') \text{ and } v'\in\mu_2^{-1}(v').$$

Thus, in order for a correspondence μ to be privacy preserving, for each $m\in M$, the set $\mu^{-1}(m)$ must be a product UxV, where U is a subset of the set $X_1=\{a,b,c\}$ and V is a subset of $X_2=\{e,f,g\}$. That is, the set $\mu^{-1}(m)$ must be a <u>rectangle</u> in $X_1 x X_2$ with side U in X_1 and side V in X_2 (c.f. Definition A1.2 and Lemma A1.2). In the case of the function F, $F^{-1}(0)$ corresponds to the entries in the (a,g) and (c,e) positions of the matrix M(F). That is, the set $F^{-1}(0)$ is {(a,g), (c,e)}. That set, {(a,g), (c,e)}, is not a product of sets $U \subseteq X_1$ and $V \subseteq X_2$. Indeed, if

$$\{(a,g),(c,e)\}=UxV,$$

then U must be the set {a,c} while V must be the set (g,e). Certainly,

$$\{a,c\}x\{g,e\} \neq \{(a,g),(c,e)\}.$$

Therefore, F is not privacy preserving.

We still face the problem of realizing F. To help in the search for a privacy preserving correspondence μ that realizes F, we examine more closely the requirement that the correspondence μ^{-1} must carry points to rectangles. Since the condition of carrying points to rectangles is a condition on the correspondence

μ^{-1}: M--->X_1xX_2, we examine correspondences from M to X_1xx_2. Suppose that we can find a correspondence ν from a set M onto the set X_1xX_2 such that ν(m) is a rectangle for each $m \in M$. Then the correspondence $\mu^{-1}:X_1xX_2$--->M is a correspondence from X_1xX_2 onto M and it certainly satisfies the requirement that $(\nu^{-1})^{-1}=\nu$ carries points of M to rectangles in X_1xX_2. we ask what other conditions $\nu^{-1}=\mu$ must satisfy in order that it be a privacy preserving correspondence. The answer is that there are no other requirements (c.f. Lemma A1.2). To see this, we construct from ν correspondences $\mu_i:X_i$--->M so that

$$\nu^{-1}(x,y)=\mu_1(x)\cap\mu_2(y).$$

The correspondence ν:M--->$X_1 x X_2$ can be composed with the projection of $X_1 x X_2$ to X_1 to produce a correspondence ν_1:M--->X_1, and similarly the composition of ν with the projection to X_2 produces a correspondence ν_2:M--->X_2. Set

$$\mu_1=\nu_1^{-1}$$

and set

$$\mu_2=\nu_2^{-1}.$$

If $m\in M$, then $\nu(m)$ is a rectangle in $X_1 x X_2$. As a rectangle,

$$\nu(m)= U \times V,$$

for some $U\subseteq X_1$ and some $V\subseteq X_2$. The projection of U x V to X_1 is U, and the projection of UxV to X_2 is V, so

$$U=\nu_1(m) \text{ and } V=\nu_2(m).$$

That is

$$\nu(m)=\nu_1(m)x\nu_2(m).$$

From this one can see that if

$$(x,y)\in X_1 x X_2,$$

then

$$\mu(x,y)=\mu_1(x)\cap\mu_2(y).$$

Indeed, if

$$n\in\mu(x,y)=\nu^{-1}(x,y),$$

then

$$(x,y)\in\nu(n)=\nu_1(n) \times \nu_2(n).$$

That is

$$x\in\nu_1(n) \text{ and } y\in\nu_2(n),$$

or what is the same thing

$$n\in\nu_1^{-1}(x)=\mu_1(x) \text{ and } n\in\nu_2^{-1}(y)=\mu_2(y).$$

Therefore,

$$\mu(x,y)\subseteq\mu_1(x)\cap\mu_2(y).$$

If, conversely,

$$n\in\mu_1(x)\cap\mu_2(y),$$

then

$$n\in\nu_1^{-1}(x) \text{ and } n\in\nu_2^{-1}(y),$$

therefore

$$(x,y)\in\nu(n),$$

and

$$n\in\nu^{-1}(x,y)=\mu(x,y).$$

Actually, this construction of the correspondences μ_i is the only way that μ can be represented as an intersection (c.f. Lemma A1.1.) More precisely, if a correspondence is privacy preserving, then the individual coordinate correspondences $\mu_i:X_i$--->M are unique.

We now have a handy way of building privacy preserving correspondences. Cover $X_1 x X_2$ with rectangles and attach a label to each of the distinct rectangles. The labels of the rectangles comprise the space M, and the correspondence ν associates each label to the rectangle with that label. The inverse correspondence, ν^{-1}, is then a privacy preserving correspondence from the set $X_1 \times X_2$ to the set of labels, M.

What conditions must a privacy preserving correspondence ν satisfy in order that ν realize our function F? Each label of a rectangle corresponds uniquely to a value of the function F. That is, each rectangle in the image of the correspondence ν is entirely contained in a level set of the function F. The function h from the set of labels M to the set $\{0,1\}$ need only assign to the label of a rectangle the value the function F assigns to the rectangle. Because the rectangle lies in one level set, the function F has the same value on every point of the rectangle.

We now build a privacy preserving correspondence ν on the set $X_1 \times X_2$ that realize F. First cover $F^{-1}(1)$ by rectangles. Refer to Figure A0.1.

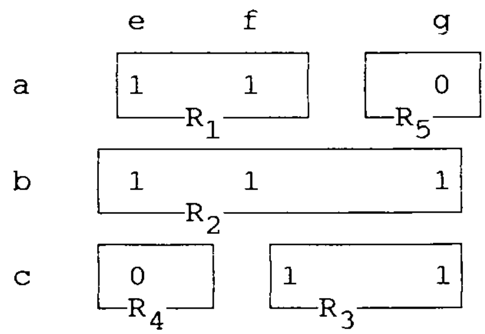

Figure A0.1

In the matrix M(F), the first row has two positions that represent points in {a,b,c}x{e,f,g} where F has the value 1. Those two points form the rectangle R_1= {a}x{e,f}. Similarly, the second row of M(F) represents the rectangle R_2={b}x{e,f,g}, and the rectangle R_2 is entirely in the level set $F^{-1}(1)$. Finally, the third row of M(F) has two 1's and those values can be covered by the rectangle R_3={c}x{e,f,g}. To complete the cover of $X_1 x X_2$={a,b,c}x{e,f,g} by rectangles, we need to cover $F^{-1}(0)$. We do this with two more rectangles, indeed there is no other choice. Set R_4={a}x{g} and set R_5={c}x{e}. The set M is the collection of labels

$$M=\{R_1,R_2,R_3,R_4,R_5\}.$$

The correspondence ν carries a label to the rectangle with that label, however, we have yet to give the correspondence $\mu=\nu^{-1}$. But this correspondence is easy to see. The point (a,e) labels an entry in the matrix

M(F) that is in the rectangle R_1. Therefore

$$\mu(a,e)=\{R_1\}.$$

Similarly,

$$\mu(c,e)=\{R_4\}.$$

The <u>outcome function</u> $h:\{R_1,R_2,R_3,R_4,R_5\}\text{--->}\{0,1\}$ carries R_1, R_2, and R_3 to 1 and R_4, R_5 to 0.

The realization $(\nu,\{R_1,R_2,R_3,R_4,R_5\},h)$ of the function F we have constructed above has the property that each point of X_1xX_2 lies in exactly one rectangle. Therefore, the correspondence ν is a function. This is certainly not a requirement. Indeed, the correspondence σ represented in Figure A0.2 is not a function, but it is a privacy preserving correspondence that realizes F. The correspondence σ carries $\{a,b,c\}x\{e,f,g\}$ to the space of labels $\{S_1,S_2,S_3,S_4\}$. The label S_1 is used for the rectangle $\{a,b\}x\{e,f\}$ and S_2 is the label for the rectangle $\{b,c\}x\{f,g\}$. As a result, $\sigma(b,f)=\{S_1, S_2\}$.

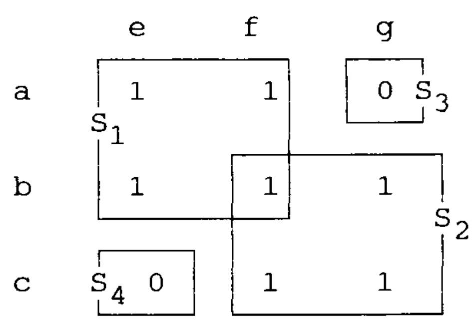

Figure A0.2

It is interesting to note that the message space for the correspondence μ has cardinality 5, and among privacy preserving correspondences that are functions 5 is the smallest cardinality possible(c.f. the discussion following Lemma A4.1 and Theorem A5.1). The privacy preserving correspondence σ has a message space of cardinality 4 and also realizes F.

The method we have outlined for building realizations relates message spaces to rectangles and allows one to check privacy in a fairly efficient way, when the sets involved have small cardinality. It is reasonable to ask if all realizations can be built by covering the level sets of a product by rectangles and then using the rectangles themselves as the elements of the message space. That is, we use as the label of a rectangle, the rectangle itself. Certainly not all privacy preserving correspondences can be characterized this way. The realizations we gave of the function F used subscripted letters for labels of the rectangles and produced message spaces that are not themselves sets of rectangles. However, it is reasonable to consider such a relabelling a realization of the function F <u>isomorphic</u> to the realization using the rectangles themselves as the members of the message space(c.f. Definition A2.2). With the concept of isomorphism in hand, we can then ask if the procedure of building privacy preserving correspondences that realize a function by covering the level sets by

rectangles produces all privacy preserving realizations of F to within isomorphism? The answer is that this procedure does not give all the privacy preserving message correspondences for realizations of a function, but the disparity is negligible (c.f. Lemma A3.2 and the discussion that follows it.). All privacy preserving correspondences arise by building rectangles, but in some cases a rectangle receives more than one label. One can then try to decide when two privacy preserving realizations are isomorphic.
The technicalities of maps between privacy preserving correspondences and the resulting concept of isomorphism are carried out in section A2.

Because {0,1} values functions on a product of two finite set X x Y can be represented easily by a matrix, as we represented the function F, above, we ask if rank conditions on the matrix yield information on the cardinality of the set of rectangles required to cover the level sets of the function F. There are some interesting bounds available given in Theorem A5.1 and Theorem A5.2.

Section A1.

We use the following notations and conventions. Most of the notation is well known, but some of the notation for various projections is non-standard.

If A is a nonempty set and if for each a in A, X_a is a nonempty set, then

$\prod_{a\in A} X_a=\prod X_a$

denotes the product of the X_a.

If $x\in\prod X_a$, then $x=\prod x_a$ where x_a is the component of x in the set X_a. For each subset S of A,

x_S

denotes the projection of x into $\prod_{a\in S} X_a$. In case S is empty x_S, is empty.

If S is a subset of A and T is a subset of $\prod X_a$ then

T_S

denotes the set

$\{\prod_{a\in S} x_a$ and $x \in T\}$.

If S is a subset of A, then

$c(S)$

denotes the complement of S in A.

If $x \in \prod_{a\in S} X_a$ and y is in $\prod_{a\in c(S)} X_a$, then

$(x\int_S y)$

is that unique element in $\prod_{a\in A} X_a$ determined by the equations

$(x\int_S y)_a=x_a$ if a is in S

$(x\int_S y)_b=y_b$, if b is in c(S).

If S is a subset of A, if U is a subset of the product $\prod_{a\in A} X_a$, and if V is a subset of the product

$\prod_{b\in c(S)} X_b$, then

$U \int_S V$

will denote the set of elements

$(u\int_S v)$ in $\prod X_a$ where $u\in U$ and $v\in V$.

If X and Y are sets, then a correspondence m from X to Y, or the graph of a correspondence from X to Y, is a subset of XxY such that for each x in X there is at least one y in Y such that (x,y) is in m. In other words, we will require that the projection into X of the graph of a correspondence from X to Y covers X. If a set m in X x Y is to be considered as the graph of a correspondence from X to Y we will write

m:X --->Y.

For each x in X we set

$m(x)=[m\cap(x \times Y)]_Y$.

That is, m(x) is the projection into X of the intersection $m\cap(x \times Y)$ where

$x \times Y=\{(x,y):y \in Y\}$.

If U is a subset of X, then

$m(U)=[m\cap(U \times Y)]_Y$.

The image of a set $Z\subseteq X$ under a correspondence m is the union of the m(z) for z in Z.

The correspondence m is onto if the image of m is Y.

We also compose correspondences. If

f:X--->Y

and

g:Y--->Z

are correspondences, then the composition of g and f, denoted by g·f, is determined by the equation

$$(g\cdot f)(x)=\cup_{y \in f(x)} g(y).$$

If F is the graph of f and G is the graph of g, this is equivalent to defining the correspondence (g·f) to have as graph the set

$$[(F \times Z)\cap(X \times G)]_X.$$

The definition of a privacy preserving correspondence is the following (c.f. [15] or [11]).

<u>Definition A1.1.</u> Assume that A is a nonempty set that indexes a collection of nonempty sets $\{X_a\}$. A correspondence

$$m:\prod_{a \in A} X_a \text{--->} M$$

is privacy preserving if for each a in A there is a correspondence

$$m_a:X_a\text{--->}M$$

such that for each x in

$$\prod X_a \quad m(x)=\cap_a m_a(x_a).$$

In case m is a privacy preserving correspondence, the correspondences $\{m_a:a\in A\}$ are called <u>coordinate correspondences</u> for the correspondence m. The space M is referred to as a <u>message space</u>.

If m is a privacy preserving correspondence from $\prod X_a$ to M, then the coordinate correspondences for m are unique. The next lemma establishes that assertion and explicitly constructs the coordinate correspondences from the graph of the privacy preserving correspondence.

Lemma A1.1. Suppose that A indexes a collection of nonempty sets $\{X_a\}$ and suppose that

$$m:\prod_{a\in A} X_a \dashrightarrow M$$

is a privacy preserving correspondence that is onto M. Set $X=\prod_{a\in A} X_a$ and denote by

$$m_a \times M$$

the projection of m, or rather the graph of m, into the set $X_a \times M$. Then,

(i) $m=\cap_a(m_{a \times M} \times X_{c(a)})$,

(ii) if $m_a:X_a \longrightarrow M$ is the correspondence with x M, then for each $x\in P$,

$$m(\ x\ )=\cap_a m_a(x_a),$$

(iii) if $m(\ x\ )=\cap_a L_a(x_a)$ for each $x\in X$ and some collection of correspondences

$$L_a:X_a \dashrightarrow M,$$

then $L_a=m_a$

for each $a\in A$.

Proof. We will build correspondences in the product $\prod_A X_a \times M$ where we use M as its own index. Because m is a correspondence from X onto M that is privacy preserving, it follows that there are correspondences

$\Gamma_a : X_a \text{--->} M$

such that for each x in X,

$m(x) = \cap_a \Gamma_a(x_a)$.

We first show that the graph of m is the intersection of the graphs

$\Gamma_a \int_{(a,M)} X_{c(a)} \subseteq \prod X_a \times M$.

If (x,t) is an element of the graph m, that is if $t \in m(x)$, then (x_a,t) is an element of Γ_a for each a. Therefore, m is contained in the intersection

$\cap_a(\Gamma_a \int_{(a,M)} X_{c(a)})$.

On the other hand, if (x,t) is an element of the intersection

$\cap_a(\Gamma_a \int_{(a,M)} X_{c(a)})$,

then for each index a, there is a y(a) in $X_{c(a)}$ such that

$((x_a,t)\int_{(a,M)} y(a))) \in (\Gamma_a \int_{(a,M)} X_{c(a)})$.

Therefore t is an element of the intersection $\cap_a \Gamma_a(x_a)$. This shows that (x,t) is in the graph of m.

We have established that when m is a privacy preserving correspondence, then the graph of m is the intersection of the graphs

$\Gamma a \int_{(a,M)} X_{c(a)}$.

What we have left to show is that this is the only way that the graph of m can be represented as an intersection of graphs of correspondences

$L_a : X_a \text{--->} M.$

We show that if for each $a \in A$, there is an $L_a : X_a \text{--->} M$ such that

$m(x) = \cap L_a(x_a)$

for all $x \in X$, then

$L_a = m_a.$

Fix an index $a \in A$. If $(z,t) \in L_a$, then for each $b \neq a$, choose a y_b in X_b so that (y_b,t) is an element of L_b. This is possible because m is assumed to be a correspondence that is onto M, and therefore for some $x \in X$, the element t is in the intersection $\cap_a L_a(x_a)$. Denote by y the element of X that has z in the a^{th} coordinate position and has y_b in the b^{th} position when $b \neq a$. The element (y,t) is in m(y) because (y,t) has been constructed as an element of $\cap_a L_a(y_a)$. The element (z,t) is the projection of the element (y,t) into the set X_a x M. Therefore

$(z,t) \in m_{a \times M}$

and hence L_a is a subset of the set $m_{a \times M}$, where $m_{a \times M}$ is the projection of the graph of m into the set X_a x M. On the other hand, if

$(w,t) \in m_{a \times M},$

then for some $x \in X$, $(x,t) \in m$ and $x_a = w$. But

$m(x) = \cap_a L_a(x_a),$

therefore

$t \in L_a(x_a)$

and

$(w,t)\in L_a$.

Therefore, $m_{a \times M}=L_a$. ▒

The principal set theoretic tool of [HRS] describes a privacy preserving correspondence in terms of the geometry of those subsets of $\prod_{a \in A} X_a$ that are the product of its projections onto the X_a.

<u>Definition A1.2.</u> Suppose that $\{X_a\}$ is a collection of nonempty sets indexed by a set A. A set T in $\prod_{a \in A} X_a$ is a <u>rectangle</u> if there are sets U_a in X_a such that $T = \prod_{a \in A} U_a$.

A useful characterization of a privacy preserving correspondence m is that when $m:\prod_a X_a$--->M is onto M, then m^{-1} transforms points into rectangles. This characterization is given in the following lemma.

<u>Lemma A1.2.</u> Suppose that $\{X_a\}$ is a collection of nonempty sets indexed by a set A and suppose that $m:\prod_{a \in A} X_a$ ---> M is a correspondence. Then m is privacy preserving if and only if for each t in M, the set $m^{-1}(t)$ is a rectangle in $\prod_{a \in A} X_a$.

Proof. Set $X=\prod_{a \in A} X_a$ and assume that m:X--->M is a privacy preserving correspondence.