DES HAS NO PER ROUND LINEAR FACTORS

J. A. Reeds and J. L. Manferdelli

AT&T Bell Laboratories
Murray Hill, New Jersey 07974

ABSTRACT

Interest in the cryptanalysis of the National Bureau of Standards' Data Encryption Standard (DES) has been strong since its announcement. Here we describe an attack on a class of ciphers like DES based on linear factors.

If DES had any non trivial factors, these factors would provide an easier attack than one based on complete enumeration. Basically, a factor of order n reduces the cost of a solution from 2^{56} to 2^n+2^{56-n}. At worst (n=1 or 55), this reduces the cost of a Diffie-Hellman search machine from 20 million dollars to 10 million dollars: a 10 million dollar savings. At best (n=28), even without iteration, the method could reduce the cost from 2^{56} to $2^{28}+2^{28}$: a computation well within the reach of a personal computer.

Alas, DES has no such linear factors.

INTRODUCTION

The basic idea here is an elaboration of a trivial idea, too good to be true. If, for each distinct value of the key, DES mapped the plaintext blocks into the ciphertext blocks *linearly*, one could deduce the matrix of that linear transformation from a small number of corresponding plaintext/ciphertext blocks. Similarly, if the dependence of the ciphertext on the key was linear, one could solve for the key. Unfortunately, the S boxes introduce strong nonlinearities: each bit output from each S box can only be represented by polynomials (in 6 variables) over GF(2) with many terms (for a discussion of these representations and their connection to coding theory see [2], chapters 2,13,14).

The current elaboration is that there might be three special linear functions of the plaintext, ciphertext and the key respectively such that the mapped ciphertext depends only on the mapped plaintext and mapped key. If the mapped key has lower dimensionality than the unmapped key, one can attempt to solve the mapped cryptosystem (possibly by brute force search).

G.R. Blakley and D. Chaum (Eds.): Advances in Cryptology - CRYPTO '84, LNCS 196, pp. 377-389, 1985.

This mapping behavior is called *cryptosystem factorization*. In general, a cryptosystem consists of a plaintext space, a key space, a ciphertext space, and a family of invertible maps indexed by the key space. We say that cryptosystem A is a *factor* of cryptosystem B if there are maps (called *factor maps*) between the plaintext, key, and ciphertext spaces such that the enciphering and deciphering actions of cryptosystem A can be recovered from those of cryptosystem B using the factor maps. If the factor mappings are linear functions we say A is a *linear factor* of B. If the key space of A is smaller than that of B one can profitably break B by first breaking A.

There is no special reason to suppose that the DES has any factors, linear or not. But if it had they probably would have the same general round-by-round flavor that DES itself has. This paper shows that the individual round of DES has no linear factors.

DES NOTATION

DES is a product cipher. The key dependent transformation that DES induces on the plaintext is a product of a family of (involutory) transformations μ and λ_i. If L and R are the two 32 bit subwords of a 64 bit input, we have*

$$\mu : LR \longmapsto RL$$

and

$$\lambda_i : LR \longmapsto L(R+f(E(L)+k_i))$$

Using these conventions,

$$DES(K,P) = IP^{-1}\lambda_{16}\mu\lambda_{15}\mu \cdots \lambda_2\mu\lambda_1 IP(LR)$$

and,

$$DES^{-1}(K,P) = IP^{-1}\lambda_1\mu\lambda_2 \cdots \lambda_{15}\mu\lambda_{16} IP(LR).$$

---

* The sign "+" in this paper denotes addition. Here, we do addition in at least three different rings: the ordinary integers [1+1=2], GF(2) [1+1=0], and vector spaces over GF(2) [(1,0,1,1)+(1,1,0,1)=(0,1,1,0)]. To emphasize that we are interested in the arithmetical properties of the "+" operator, we use + in all three cases. We rely on the reader to distinguish which ring (and hence which operator) is being used in any given equation. In the second displayed equation, for example, the first plus denotes addition done in the vector space of dimension 32 over GF(2); the second plus refers to arithmetic done in the vector space of dimension 48 over GF(2). Readers who are not familiar with DES will see in a few paragraphs why the rings in the second equation are what we say they are.

The transformation IP consists of a permutation of the input bits; it has no cryptographic significance and need not be mentioned any further. E, P and the S boxes S_1 ,..., S_8 are defined in [1] and will be discussed in more detail below. k_i (i=1,2,3,...,16) is a 48 bit subkey for round i derived from a 56 bit key **k** according to a key schedule described in [1]. We refer to the composed map $\sigma_i = \mu\lambda_i$ as a "round" of DES. Note that DES is composed of 16 encrypting rounds with the switch of the 32 bit subwords suppressed in the final round.

Denoting the vector space of dimension n over GF(2) by V_n, we have:

$$E\colon\ V_{32} \longrightarrow V_{48}$$

$$P\colon\ V_{32} \longrightarrow V_{32}$$

$$f\colon\ V_{48} \longrightarrow V_{32}$$

$$S_i\colon\ V_6 \longrightarrow V_4$$

E is the expansion matrix which takes $x = (x_1, \ldots, x_{32})$ to $(x_{E(1)}, \ldots, x_{E(48)})$. The function f is obtained by applying successive S boxes to the successive six bits of the argument and then applying the permutation matrix P to the resultant vector, *i.e.*:

$$y = (S_1(x_1, \ldots, x_6), \ldots, S_8(x_{43}, \ldots, x_{48}))$$

$$f(x) = (y_{P(1)}, \ldots, y_{P(32)})$$

E and P are linear functions, f is not. Writing this in tabular form, we get

Two rounds of DES		
round	left 32 bits	right 32 bits
0	L	R
1	R	$L+f(E(R)+k_1)$
2	$L+f(E(R)+k_1)$	$R+f(E[L+f(E(R)+k_1)]+k_2)$

It is convenient to employ another set of equations to describe DES. Setting $x_0 = E(L)$, $x_1 = E(R)$ and $x_2 = E(L+f((E(R)+k_1))$, we can write a recurrence based on the second column of the table above.

$$x_0 + x_2 = \phi(x_1+k_1)$$

where

$$\phi(x) = Ef(x) \tag{1}$$

In fact, if we write down all 16 rounds (and perform an extra switch of the two 32 bit subwords at the end), we see that

$$x_{i+1} + x_{i-1} = \phi(x_i + k_i) \tag{2}$$

given the obvious definition for x_i for $i = 1,2,...,16$. With this notation, the output of the DES algorithm consists of two 32 bit subwords of x_{16} and x_{15}.

PER ROUND LINEAR FACTOR OF TYPE A

For reasons that will become clear momentarily, we would like to find a matrix A, and a function ψ such that

$$A\phi(x) = \psi(Ax). \tag{3}$$

for all x. Under these conditions, we say we have an *A factor*, in honor of A occurring in equation (3) above. If (1), (2), and (3) hold,

$$Ax_{i+1} + Ax_{i-1} = \psi(Ax_i + Ak_i)$$

yielding

$$y_{i+1} + y_{i-1} = \psi(y_i + l_i) \tag{4}$$

where $y_i = Ax_i$, $l_i = Ak_i$. Equation (4) is identical in form to equation (2), so the pairs (y_i,l_i) form a new cipher system. We call this the "mapped" cipher system. y_0,y_1 form the mapped plaintext, y_{15} and y_{16} form the mapped ciphertext and the l_i are the mapped per round keys.

Let KS_i be the key schedule matrix for round i, then the map

$$\mathbf{k} \longmapsto (KS_1(\mathbf{k}),...,KS_{16}(\mathbf{k}))$$

has an image (in V_{768}) of dimension 56. If the corresponding key schedule for the mapped cipher, given by

$$(A\ KS_1(\mathbf{k}),...,A\ KS_{16}(\mathbf{k})) = \mathbf{l} = (l_1,...,l_{16}),$$

has dimension n $(0 < n < 56)$, we can recover the original key as follows. Search over the mapped keyspace to find the $\mathbf{l}$ producing the correct behavior in a transformed plain/ciphertext pair. This costs 2^n time. Then go back to the original cipher, looking for the key $\mathbf{k}$ in the coset of the null space of A mapping to $\mathbf{l}$. This costs 2^{56-n} time. Total cost: $2^n + 2^{56-n}$.

We need some more notation for later, most of the notation concerns projection operators of various sorts to wit:

$\pi_i(x_1,\ldots,x_n) = (0,0,\ldots,x_i,0,\ldots,0)$

$\theta_{j,m} = \pi_j + \cdots + \pi_m$

$\rho_i = \theta_{6(i-1)+1,6i}$

$\rho'_i = \theta_{4(i-1)+1,4i}$

$V^{(i)} = \rho_i(V_{48})$

$V'^{(i)} = \rho'_i(V_{32})$

$W^{(i)} = E(V'^{(i)})$

$\phi_x(y) = \phi(x+y) - \phi(x)$

N_A is the null space of A.

LOOKING FOR AN A

Now we show that no such non trivial A exists. The following characterization will facilitate the search for A. Statement 1 is the one we want for cryptanalysis. Statement 2 is easier to verify; statement 3 is still easier to verify.

THEOREM 1. Suppose $A:V \longrightarrow V$ and $\phi:V \longrightarrow V$, with A linear. The following are equivalent.

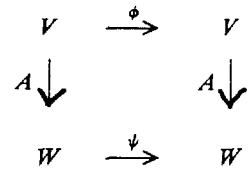
1. There is a $\psi:W \longrightarrow W$ such that $A\phi(x) = \psi(Ax)$.
2. If $Ax = Ay$ then $A\phi(x) = A\phi(y)$.
3. For all x in V, $\phi_x(N_A) \subseteq N_A$.

PROOF. 3 $\Longrightarrow$ 2: If $Ax = Ay$, $A(x-y) = 0$ so $x-y$ is in N_A. By the conclusion of 3, $A(\phi(z+(x-y)) - \phi(z)) = 0$ for all z in V. Setting $z = y$ and distributing the A, we get $A\phi(x) = A\phi(y)$.

2 $\Longrightarrow$ 3: If z is in N_A, $A(x+z) = A(x)$ for any x; so, by 2, $A\phi(x+z) = A\phi(x)$. Thus, $A(\phi(x+z) - \phi(x)) = 0$; so $\phi(x+z) - \phi(x)$ is in N_A.

1 $\Longrightarrow$ 2: $A(\phi(x) - \phi(y)) = \psi(Ax) - \psi(Ay) = 0$, the last equality follows from 1 if $x = y$.

2 $\Longrightarrow$ 1: Define $\psi(Ax) = A\phi(x)$. We need only show that the given map is well defined. If $Ax = Ay$, $A\phi(x) = A\phi(y)$, so the map is well defined. Note that W is just the image (in V) of A. This condition insures that the diagram below commutes.



Commuting diagram for 2 $\Rightarrow$ 1

By Theorem 1 (3), we want to look for subspaces S satisfying the following condition.

CONDITION S. $\phi_x(S) \subseteq S$ for all x in V.

THEOREM 2. Let $T_i(a) = \text{span}\{ S_i(a+b)-S_i(b)$, all b in $V_6 \}$. If $i \neq 4$ and $a \neq 0$ then $T_i(a)$ is $V^{(i)}$. If $i=4$ and $a \neq 0$, $T_i(a)$ is one of two 2 dimensional spaces, a 3 dimensional space or the entire 4 dimensional space, $V^{(i)}$.

PROOF. A simple computer program was written to verify these.

THEOREM 3. Suppose S is a subspace satisfying "condition S". Further, suppose there is a y in S with $\rho_i(y) \neq 0$. If $i \neq 4$, $W^{(i)} \subseteq S$; if $i=4$, S contains at least a two dimensional subspace of $W^{(i)}$.

PROOF. Suppose u,v are in $V^{(i)}$. Then by condition S,

$$u^* = \phi(y+u)-\phi(u) = EP(S_1(\rho_1(u+y))-S_1(\rho_1(u)), \ldots, S_8(\rho_i(u+y))-S_8(\rho_i(u)))$$

and

$$v^* = \phi(y+v)-\phi(v) = EP(S_1(\rho_1(v+y))-S_1(\rho_1(v)), \ldots, S_8(\rho_i(v+y))-S_8(\rho_i(v)))$$

are in S. u^*-v^* must also be in S and $\rho_j(u+y) = \rho_j(v+y) = \rho_j(y)$ if $j \neq i$. So,

$$u^*-v^* = EP(0,0,\ldots,0,S_i(\rho_i(u+y))-S_i(\rho_i(u))-S_i(\rho_i(v+y))+S_i(\rho_i(v)),0,\ldots,0)$$

is also in S. Setting

$$T(u,v) = S_i(\rho_i(u+y))-S_i(\rho_i(u))-S_i(\rho_i(v+y))+S_i(\rho_i(v)),$$

theorem 2 tells us that $\text{span}\{T(u,v)\colon u,v \in V^{(i)}\}$ is all $V^{(i)}$ if $i \neq 4$ and is at least a two dimensional subspace of $V^{(i)}$ if $i=4$. Thus, if $i \neq 4$ S contains $EP(V^{(i)}) = W^{(i)}$; if $i=4$, $EP(S)$ is (at least) a two dimensional subspace of $W^{(i)}$. QED.

REMARK. We say output block k is affected by input block i if at least one of the bits of $V^{(k)}$ is calculated using S box i. EP switches and expands outputs from the S box calculation so its easy to see that output block k is affected by input block i iff $V^{(k)} \cap W^{(i)} \neq 0$.

Effect of 6 bit (input, output) blocks on (x, $\phi(x)$)		
In block	Out block (round 1)	Out block (round 2)
1	7,4,2,5,6,8	all blocks
2	6,8,3,7,5,1	all blocks
3	5,1,4,6,7,2	all blocks
4	7,2,5,8,3,1	all blocks
5	3,1,2,6,4,8	all blocks
6	4,8,7,1,3,5	all blocks
7	3,5,4,2,8,6	all blocks
8	2,6,3,1,7,4	all blocks

LEMMA. Suppose S is a subspace satisfying "condition S" and suppose $i \neq 4$. If a 6 bit output block k is affected, during the calculation of $\phi(x)$, by a bit from a six bit input block i and if S contains a y, such that $\rho_i(y) \neq 0$, then $W^{(k)} \subseteq S$ provided $k \neq 4$. If $k=4$, there is at least a two dimensional subspace of $W^{(k)}$ contained in S.

PROOF. If output block k is affected by input block i, $\rho_k(W^{(i)}) \cap V^{(k)} \neq 0$. Since $\rho_i(y) \neq 0$, theorem 3 yields $W^{(i)} \subseteq S$; this, in turn, means there is a y in S such that $\rho_k(y) \neq 0$. Applying theorem 3 again, we get $W^{(k)} \subseteq S$, if $k \neq 4$; if $k=4$ there is a two dimensional subspace, W', $W' \subseteq S$, with $W' \subseteq W^{(k)}$. This is exactly what the lemma claims, so we are done. QED.

THEOREM 4. If S is a subspace satisfying "condition S" and $S \neq 0$ then $S = W$.

PROOF. We prove this by pumping up S to W. Suppose $S \neq 0$, then there is an i $(1 \leqslant i \leqslant 8)$ and a y in S with $\rho_i(y) \neq 0$.

For the sake of simplicity, let's assume $i=1$, so $\rho_1(y) \neq 0$. By theorem 1, $W^{(1)} \subseteq S$; by the Lemma, $W^{(k)} \subseteq S$, for $k = 2,5,6,7,8$ and, in addition, there's at least a 2 dimensional subspace of $W^{(4)}$ in S. Now, $W^{(7)} \subseteq S$ so by reapplying the Lemma, we get $W^{(k)} \subseteq S$, for $k = 2,3,5,6,8$. To recap, $\rho_1(y) \neq 0$ implies that $W^{(k)} \subseteq S$ for $k \neq 4$.

$\sum_{k=1,3,5,6,8} (W^{(k)} \cap V^{(4)}) = V^{(4)}$. It's easy to see that $V^{(4)} \subseteq S$ implies $W^{(4)} \subseteq S$. Thus $W^{(k)} \subset S$ for $1 \leqslant k \leqslant 8$, hence $S = W$.

For values of i other than 1 and 4 the argument in the preceding paragraph applies *mutatis mutandis*. If $\rho_4(y) \neq 0$, the table and Theorem 3 show that $W^{(4)} \cap V^{(n)} \neq 0$ for some n in $\{1,2,3,5,7,8\}$. Thus, for some y in S $\rho_i(y) \neq 0$, for some $i \neq 4$ provided only that $S \neq 0$. By the above argument, $S = W$. QED.

REMARK. The proof of the theorem above basically "reapplies" the mapping $\phi_x(y)$ for non zero y in S until ϕ_x gobbles up S.

EXTENSION TO AB FACTORS

We have called the sort of per round linear factor discussed above an *A factor*, in honor of the equation

$$A\phi(x) = \psi(Ax)$$

which holds for all x. A fancier kind of factor is the *AB factor*, which we now discuss. Here we suppose we are given a pair of linear maps A and B, and a possibly non linear function ψ, so that for all x both of

$$A\phi(x) = \psi(Bx)$$

and

$$B\phi(x) = \psi(Ax)$$

hold. Clearly an A factor is an AB factor: just let $B = A$.

A non trivial AB factor can also be used to solve the DES. One applies A and B alternately to the DES rounds. Let

$$T_i = A$$

if i is even and

$$T_i = B$$

if i is odd. Then

$$y_i = T_i\ x_i$$

$$l_i = T_i\ k_i$$

and

$$y_{i+1} + y_{i-1} = \psi(y_i + l_i)$$

as in equation (4) above. This is a factor cryptosystem of DES type, but it may have a smaller keyspace.

Unfortunately, if an AB factor exists, so does an A factor. This follows from the following fact, whose proof is easy:

THEOREM 5. Let

$$\phi\colon V \longrightarrow V$$

$$\psi_1\colon W_1 \longrightarrow W_2$$

$$\psi_2\colon W_2 \longrightarrow W_1$$

$$T_1\colon V \longrightarrow W_1$$

$$T_2\colon V \longrightarrow W_2$$

be maps between vector spaces, ϕ, ψ_1, and ψ_2 not necessarily being linear. Suppose, for all v in V we have

$$T_2(\phi(v)) = \psi_1(T_1(x))$$

and

$$T_1(\phi(v)) = \psi_2(T_1(x)).$$

Then there is a vector space W and a linear map $A\colon V \longrightarrow W$ and a function $\psi\colon W \longrightarrow W$ such that for all v in V,

$$A(\phi(v)) = \psi(A(x)).$$

PROOF. Let $W = W_1 \oplus W_2$. Then $A(v) = (T_1(v), T_2(v))$ and $\psi(w_1, w_2) = (\psi_1(w_1), \psi_2(w_2))$ satisfy the conclusion of the theorem. QED.

EXTENSION TO $\alpha\beta\gamma$ FACTORS

Stepping back a moment, we might say that the point of the above attack is to find a per round *linear* relationship among the (*plaintext*,*ciphertext*,*key*) triples. If we don't insist that the relationship be linear, a broader attack may hold. For example, consider the following, which we call an $\alpha\beta\gamma$ factorization.

Let σ_i be a basic enciphering operation (like a round of DES) depending on the key bits $\mathbf{k}_i$. Then the plaintext $\mathbf{x}_0$ is converted to the ciphertext $\mathbf{x}_n$ by the iteration $\mathbf{x}_{i+1} = \sigma_i(\mathbf{x}_i)$. Suppose we can find α, β, and γ with γ linear satisfying

$$\alpha(\mathbf{x})+\beta(\sigma_i(\mathbf{x}))+\gamma(\mathbf{k}_i) = 0 \tag{5}$$

and

$$\beta(\mathbf{x}) + \alpha(\sigma_i(\mathbf{x})) + \gamma(\mathbf{k}_i) = 0. \tag{6}$$

(Equivalently, we might require

$$\alpha(\mathbf{x})+\alpha(\sigma_i(\mathbf{x})) = \beta(\sigma_i(\mathbf{x}))+\beta(\mathbf{x}) \tag{7}$$

instead of (6).) Now we can apply (5) to the enciphering equations, yielding (term by term)

$$\alpha(\mathbf{x}_i)+\beta(\mathbf{x}_{i+1}) = \gamma(\mathbf{k}_i)$$

and, on summation,

$$\sum_{i=0}^{n-1}\alpha(\mathbf{x}_i) + \sum_{i=0}^{n-1}\beta(\mathbf{x}_{i+1}) = \sum_{i=0}^{n-1}\gamma(\mathbf{k}_i).$$

Rearranging and using (2) and canceling terms appearing an even number of times, we get

$$\alpha(\mathbf{x}_0) + \alpha(\mathbf{x}_n) = \sum_{i=0}^{n-1}\gamma(\mathbf{k}_i) \tag{7}$$

when n is even and

$$\beta(\mathbf{x}_0) + \alpha(\mathbf{x}_n) = \sum_{i=0}^{n-1}\gamma(\mathbf{k}_i)$$

when n is odd. Belaboring the point, we might write

$$\alpha(\textit{plaintext}) + \alpha(\textit{ciphertext}) = \sum_{i=0}^{n-1}\gamma(\mathbf{k}_i) \tag{8a}$$

or

$$\beta(\textit{plaintext}) + \alpha(\textit{ciphertext}) = \sum_{i=0}^{n-1}\gamma(\mathbf{k}_i). \tag{8b}$$

To use such a relation to help find the key, suppose we are trying to find a key in V_k. Let W be the $k-1$ dimensional subspace satisfying (8a) or (8b). Instead of searching all elements in V_k, restrict the search to elements of W. This produces a computational savings of ½; if many such relations can be found, determination of the key (even for a large keyspace) would be quick and painless.

Whether α,β,γ satisfying (5) and (6) exist is a deep question. Sometimes their existence and discovery are not too difficult. For example, Equation (6) automatically holds if (5) holds and σ_i is an involution. Significantly, it is easy to show that (6) also holds for a round of DES (where $\sigma_i=\mu\lambda_i$) if (5) holds with $\sigma_i = \lambda_i$ and $\alpha(\mu(\mathbf{x}))+\beta(\mu(\mathbf{x})) = \alpha(\mathbf{x})+\beta(\mathbf{x})$. If an S box had a non trivial affine dependence, we could manufacture such functions in the following manner. Suppose we had

$$MS(\mathbf{x}) + L(\mathbf{x}) = 0$$

for some S box S with M and L matrices of size 1×4. Set $\alpha = M$ and $\beta = \gamma = L$. As a consequence of the above relation, we have

$$\alpha(S(\mathbf{x}_i)) + \beta(\mathbf{x}_i) = 0$$

or since $\beta=\gamma$ is linear

$$\alpha(S(\mathbf{x}_i)) + \beta(\mathbf{x}) = \gamma(\mathbf{k}_i)$$

It is easy to see how to modify α,β and γ when we replace S by the σ_i of DES.

Once again, no S box has this linear property. But this begs the larger question: *Do any such α,β,γ exist for DES?* Unfortunately, it can be shown that any $\alpha\beta\gamma$ factor already takes this form. To see this, it will be convenient to switch notation. Writing a round of DES as

$$(\mathbf{x},\mathbf{y}) \longmapsto (\mathbf{y},\mathbf{x}+f(\mathbf{y}+\mathbf{k}_i))$$

equation (5) becomes

$$\alpha(\mathbf{x},\mathbf{y}) + \beta((\mathbf{y},\mathbf{x}+f(\mathbf{y}+\mathbf{k}_i))) + \gamma(\mathbf{k}_i) = 0 \tag{9}$$

f, being the cryptographic function defined in the section on DES notation. Now set $\mathbf{y} = 0$, and, as before, let $\phi(x) = Ef(x)$. Let P be a quasi inverse of E on V_{32}, i.e., $PE(x) = x$ for x in V_{32}. Now make the following definitions (Caution: the f below is not the same as the f in equation (5); also, remember that P is *not* the P defined in the section on DES notation.)

$$g(x) = \alpha(Px,0)$$

$$h(x) = \gamma(x)$$

$$f(x) = \beta(0,Px)$$

Then for all x and y in W $(=V_{48})$,

$$g(x) + h(y) = f(x+\phi(y))$$

We now show that the above equation holds only if f is affine.

THEOREM 6. Suppose $\phi: W \longrightarrow W$ and that for all x, y in W

$$f(x+\phi(y)) = g(x) + h(y)$$

where h is linear and $g(a) = 0$ for some a in W. If Image(ϕ) is an abelian group then f is affine on Image(ϕ). Since f is affine and h is linear, g is also affine.

PROOF.

$$f(x+\phi(y_1+y_2)) = g(x)+h(y_1)+h(y_2)=f(x+\phi(y_1))+h(y_2) = f(x+\phi(y_2))+h(y_1) \qquad (*)$$

Since we are in GF(2),

$$f(x+\phi(y_1)) = f(x+\phi(y_2)) + h(y_1) + h(y_2)$$

$g(a) = 0$ and (*) imply

$$f(x+\phi(y_1)) = f(x+\phi(y_2)) + f(a+\phi(y_1)) + f(a+\phi(y_2))$$

Since the image of ϕ is in W and a is in W, we can set $x = a + \phi(y_2)$ yielding

$$f(\phi(y_1)+\phi(y_2)+a) = f(a+\phi(y_1))+f(a+\phi(y_2))+f(a)$$

Finally, since Image(ϕ) is an abelian group, for all u_1, u_2 in Image(ϕ), we can find y_1, y_2 in W with $u_1 = \phi(y_1)$, $u_2 = \phi(y_2)$ giving:

$$f(u_1+u_2+a) = f(a+u_1) + f(a+u_2) + c.$$

Setting $l(x) = f(x+a)$, this becomes

$$l(u_1+u_2) = l(u_1) + l(u_2) + c$$

as claimed.

THEOREM 7. DES has no per round linear factors of α,β,γ type.

PROOF. If DES had a per round factor of α,β,γ type, then by theorem 6, the factor functions would express a (non trivial) affine relationship among the input, output, and key bits of a round. Since the outputs of different S boxes are algebraically independent, it suffices to show that no such relationship holds among the four output bits of any of the eight S boxes. Application of the following lemma concludes the proof.

LEMMA. Let $S_{4(i-1)+j}$ denote the j'th bit of S box i. Then for all i,

$$\sum_{j=1}^{4} a_j S_{4(i-1)+j}(x_1, \ldots, x_6) + \sum_{j=1}^{6} b_j x_j + d = 0$$

implies that $a_j = b_k = 0$ for $j = 1,2,3,4$, $k = 1,2,\ldots,6$.

PROOF. Linear algebra applied to the truth tables of all of the output bits of all of the S boxes. QED.

CONCLUSION

DES seems to have no non trivial linear per round factor structure. It's hard to imagine a non linear per round factor structure that is useful for cryptanalysis. It is barely possible DES has a non trivial global factor structure that induces trivial factor behavior per round but nobody we know has a clue about what that would look like. The conclusion is that DES will not be solvable by factorization.

Nothing in this note says anything about *approximate* factorizations, or factorizations that *usually* hold, nor have we given up on finding non linear per round factors that yield tractable (non linear) constraint equations.

References

[1] National Bureau of Standards, Data Encryption Standard. U.S. Department of Commerce, FIPS **46**, 15 January 1977.

[2] MacWilliams, F.J. and N.J.A. Sloane, The Theory of Error Correcting Codes. North-Holland, 1977.