FINDING f-FREE SUBSETS OF MAXIMAL CARDINALITY

PIERRE BOUCHARD AND YEONG-NAN YEH

31 March 1992

ABSTRACT. Given a (not necessarily everywhere defined) endofunction f on a set T, we are interested in f-free subsets of T, that is subsets S such that $s \in S$ and $f(s)$ defined imply $f(s) \notin S$. We give an algorithm for explicitly finding such a subset of maximal cardinality. We then compute such cardinality in several classes of examples.

1. HISTORY.

The problem of finding *free* subsets goes back to early trials towards Fermat's conjecture when people were hoping to find a counter-example modulo p. They would study *sum-free* subsets of Z_p, that is subsets S such that the sum of two elements of S is no more an element of S. Then came 2-*free* subsets, that is those subsets S for which $x \in S$ implies $x + x \notin S$. These were studied in many contexts: notably in near-rings [4], [8], in Z_n, and in the dihedral group D_n. In the last two groups, Yeh [9] also studied q-free subsets and found their maximal possible cardinality. In the present paper we generalise the functions (like $x \mapsto qx$), with respect to which S is defined to be free, to include *any* function f in *any* finite set T and we provide an algorithm for finding an f-free subset S of T of maximal cardinality.

2. INTRODUCTION.

Given a set T and a function $f\colon T_0 \to T$ where T_0 is a subset of T, we shall call a subset $S \subseteq T$ an *f-free* subset if and only if

$$x \in S \cap T_0 \Longrightarrow f(x) \notin S,$$

in other words, if $x \in S$ and $f(x)$ is defined, then $f(x) \notin S$.

A function $f\colon T_0 \to T$ with $T_0 \subseteq T$ will be called a *subendomorphism* of T. The subendomorphisms can be made into a species by setting, for any finite set U and any bijection $g\colon U \to V$:

$$\begin{aligned} \mathrm{SubEnd}[U] &= \{\, f \mid f\colon U_0 \to U \text{ and } U_0 \subseteq U \,\} \\ \mathrm{SubEnd}[g] &= gfg^{-1}|_{f(U_0)}. \end{aligned}$$

By abuse of language, we shall often refer to the subendomorphism (T, f) in order to stress the set T and to a subset of (T, f) in order to to stress the function f.

For example, if $T = [5] = \{1, 2, 3, 4, 5\}$ and $f(x) = 2x$, then $f(x)$ is defined only for 1 and 2 and $S = \{2, 3, 5\}$ is f-free.

A subset S of a subendomorphism (T, f) will be called *max-f-free* if it is a f-free subset of T having the maximum possible cardinality. The set S of the above example is not max-f-free although it is maximal among the f-free subsets of T. This example shows that the f-free subsets of T do not always define a matroid.

The maximum cardinality of an f-free subset of a subendomorphism (T, f) will be denoted $\psi(T, f)$ or just $\psi(T)$ when this does not lead to confusion.

Let us recall that a digraph is *functionnal* if each of its vertices has at most one predecessor, that is if it is the digraph $G_f = (V, E)$ of a function $f\colon A \to B$ where $V = A \cup B$ and $E = \{(a, f(a)) \,|\, a \in A\}$. There is a one-to-one correspondence between functionnal digraphs and subendomorphisms. We shall therefore sometimes refer to $\psi(G)$ for a functional digraph, meaning $\psi(T, f)$ where f is the function such that $G = G_f$.

We can now reformulate the definition of an f-free subset in terms of the digraph G_f of f.

Lemma 2.1. *A set S is a f-free subset of a subendomorphism (T, f) if and only if no edge of G_f has both its endpoints in S.*

proof: obvious.

We shall also make extensive use of the next lemma.

Lemma 2.2. *If G_f is made of k connected components $G_1, \dots\ G_k$, then*

$$\psi(G_f) = \psi(G_1) + \cdots + \psi(G_k).$$

proof: Follows from the preceeding lemma.

3. Algorithm

In the present section, we present the algorithm for finding a max-f-free subset of a subendomorphism (T, f) and illustrate it with some of the examples that we have computed.

Let us recall that a *leaf* in a digraph $G = (V, E)$ is a vertex without predecessor. We shall denote by $L(G)$ the set of leaves of G.

Given a subset W of the set V of vertices of the digraph G, the notation $G \setminus W$ shall mean the digraph

$$G \setminus W = (\, V \setminus W,\ E \setminus \{\, (x, y) \,|\, \{x, y\} \cap W \neq \varnothing \}\,)$$

obtained from G by removing the vertices in W and all edges connecting to those vertices. By abuse of language, $G = \varnothing$ means $G = (\varnothing, \varnothing)$.

Algorithm.

INPUT: A function $f\colon T_0 \to T$ where $T_0 \subseteq T$.
OUTPUT: A max-f-free subset S of (T, f).

```
begin
    V := T;
    E := { (a, f(a)) | a ∈ T_0 };
    G := (V, E);                          we start with G = G_f
    S := ∅;
    F := L(G);
    while F ≠ ∅ do                        at the end of this while loop
        pick x ∈ F;                       G will be a disjoint union of
        S := S ∪ {x};                     directed cycles
        if x ∈ T_0 then G := G \ {x, f(x)}
                   else G := G \ {x}
                   endif;
        endwhile;
    for C ∈ { D | D is a cycle of G } do
        pick x ∈ C;
        C := C \ {x};                     this makes C a directed line
        while C ≠ ∅ do
            pick y ∈ L(C);                y is the beginning of line C
            S := S ∪ {y};
            if y ∈ T_0 then C := C \ {y, f(y)}    we remove y and the next vertex
                       else C := C \ {y}          on the line C if it exists
                       endif;
            endwhile;
        endfor;
    RETURN(S);
    end.
```

Some remarks are appropriate:

(1) The algorithm allows us to find in at most $|T|$ steps, a max-f-free subset of T, making computations of max-f-free subsets of a T as big as $\mathfrak{S}_7$ feasable in a few minutes without the help of a computer. (see example 3)
(2) The only choices we have are in the "*pick* $x \in F$" and the "*pick* $x \in C$. These may lead to many different max-f-free subsets S but not necessarily all of them. (see example 1)

Example 1. Illustration of the algorithm.

The next sequence of pictures shows what might happen to the digraph G_f when it is acted upon by the algorithm.

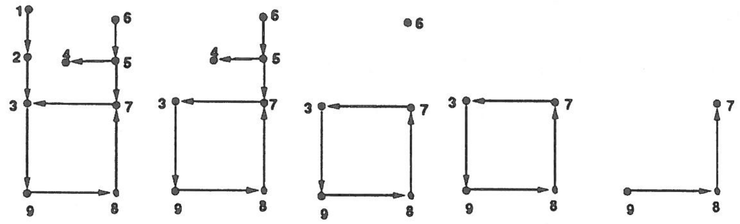

Figure 1. Illustration of the algorithm

We see that during the excution of the algorithm, S here takes in succession the values $S = \varnothing$, $\{1\}$, $\{1,4\}$, $\{1,4,6\}$, $\{1,4,6,7\}$ and $\{1,4,6,7,9\}$ which is a max-f-free subset of $1,\dots,9$.

Had we chosen to remove 7 or 9 from the cycle, we would have obtained the max-f-free subset $S = \{1,3,4,6,8\}$ instead.

Notice however that the max-f-free subset $\{2,4,6,7,9\}$ could not have been obtained by the algorithm. Notice also that $\{1,3,5,8\}$ is a maximal f-free subset of [9] but not a max-f-free subset because it has only four elements.

In the next 3 examples, we shall study $f(x) = x^2$ in different contexts: finite fields, integers modulo an odd prime power and symmetric groups.

Example 2. The finite field F_q, $q = p^n$, p prime.

This example gives us an idea of the form that may take the cardinality of a max-f-free subset.

Here the digraph G_f has a component $\{0\}$ (with a loop at 0) and on the complement of $\{0\}$ is like the digraph of the integers modulo $q-1$, the multiplicative group of F_q being cyclic. The max-f-free subsets of F_q^* are therefore in bijection with the max-2-free subsets of $\mathbb{Z}_{q-1}$, (that is max-$(x \mapsto 2x)$-free subsets of $\mathbb{Z}_{q-1}$). Therefore

$$\psi(F_q, x \mapsto x^2) = \psi(\mathbb{Z}_{q-1}, x \mapsto 2x).$$

The last term of the above equation has been computed in [9] and this gives us the following proposition.

Proposition 3.1. *The cardinality of a max-f-free subset of the finite field F_q, where $f(x) = x^2$ and $q - 1 = 2^m p_1^{e_1} \cdots p_k^{e_k} = 2^m \boldsymbol{p}^{\boldsymbol{e}}$ is equal to*

$$\boldsymbol{p}^{\boldsymbol{e}} \cdot \sum_{0 \le j \le \lfloor \frac{m-1}{2} \rfloor} 2^{m-1-2j} + \chi(m \text{ is even}) \cdot \sum_{0 < \boldsymbol{a} \le \boldsymbol{e}} \frac{\phi(\boldsymbol{p}^{\boldsymbol{a}})}{d_{\boldsymbol{a}}} \lfloor \frac{d_{\boldsymbol{a}}}{2} \rfloor$$

where $d_{\boldsymbol{a}} = \mathrm{ord}_{\boldsymbol{p}^{\boldsymbol{a}}}(2)$ and $\boldsymbol{a} = (a_1, \dots a_k) \le (e_1, \dots e_k) = \boldsymbol{e}$ means $a_i \le e_i$ for $1 \le i \le k$, ϕ is Euler's ϕ-function and $\chi(P) = 1$ if P is true, $\chi(P) = 0$ otherwise.

With the help of the computer algebra program Maple, we have been able to compute some values of $\psi(F_{p^\alpha})$ for small values of p and α:

alpha	p=2	p=3	p=5	p=7
1	0	1	2	3
2	1	5	15	31
3	2	13	74	171
4	7	52	409	1 575
5	12	121	1 951	8 403
6	30	455	9 765	77 206
7	54	1 093	48 827	411 771
8	127	4 305	256 347	3 828 187
9	226	9 841	1 220 699	20 176 803
10	508	36 905	6 103 515	185 374 380

Example 3. The multiplicative monoid $\mathbb{Z}_{p^\alpha}$.

Here the trick is to observe that $\mathbb{Z}_{p^\alpha}$ is the disjoint union

$$\mathbb{Z}_{p^\alpha} = D(\mathbb{Z}_{p^\alpha}) \cup U(\mathbb{Z}_{p^\alpha})$$

where $D(\mathbb{Z}_{p^\alpha})$ is the set of zero-divisors of $\mathbb{Z}_{p^\alpha}$ and $U(\mathbb{Z}_{p^\alpha})$ is the set of units of $\mathbb{Z}_{p^\alpha}$. We know ([6], th. 2.25) that $U(\mathbb{Z}_{p^\alpha})$ is a cyclic group of order $\phi(p^\alpha) = p^{\alpha-1}(p-1)$ and therefore

$$\psi(U(\mathbb{Z}_{p^\alpha}), x \mapsto x^2) = \psi(\mathbb{Z}_{\phi(p^\alpha)}, x \mapsto 2x)$$

which is known from [9]. The functionnal graph of $(x \mapsto x^2)$ on $D(\mathbb{Z}_{p^\alpha})$ is a graph implosing on 0 to which the algorithm is easily applied. For example, if $p = 5$ and $\alpha = 2$ we get

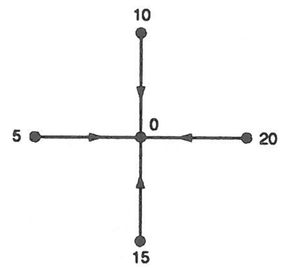

Figure 2. $D(\mathbb{Z}_{25})$

and a max-$(x \mapsto x^2)$-free subset has 4 elements and if $\alpha = 3$ we get

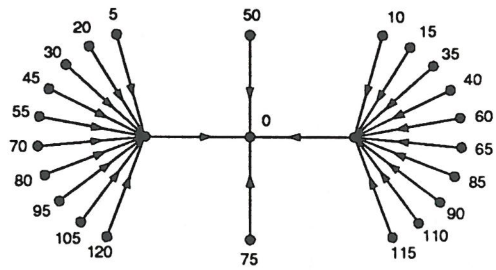

Figure 3. $D(\mathbb{Z}_{125})$

and a max-($x \mapsto x^2$)-free subset has 22 elements. The case $p = 5$ and $\alpha \in \{1,2,3,4,5\}$ is summarized in the next table:

α	5^α	$\psi(U(\mathbb{Z}_{5^\alpha}))$	$\psi(D(\mathbb{Z}_{5^\alpha}))$	$\psi(\mathbb{Z}_{5^\alpha})$
1	5	2	0	2
2	25	12	4	16
3	125	62	22	84
4	625	312	114	426
5	3 125	1 562	552	2 114

Example 4. The symmetric group $\mathfrak{S}_n$.

Here we observe that the type of the square of a permutation depends only on the type of the permutation itself. That gives us immediately the shapes of the connected components of G_f, looking only at the partititons of n, that is at the type of permutations of n. All the vertices of such a connected component eventually lead to one cycle of length k, (possibly $k = 1$). The permutations inside such a cycle will all be of the same type. Therefore we know that the number of connected components of a given shape containing a cycle of lenght k made of permutations of type $1^{a_1}2^{a^2}\cdots n^{a_n}$ is

$$\frac{1}{k}\,\frac{n!}{1^{a_1}a_1!\,2^{a_2}a_2!\cdots n^{a_n}a_n!}.$$

We can compute a max-f-free subset for each such shape using our algorithm and then apply our Lemma 2.2.

Let us illustrate this for the case of $\mathfrak{S}_6$. There are four shapes summarized in the following diagrams (fig. 4), where, for example, the number 3 on the edge from the vertex $1\,2\,3$ to the vertex $1^3\,3$ means that a given permutation of type $1^3\,3$ is the square of 3 different permutations of type $1\,2\,3$. This avoids otherwise messy digraphs.

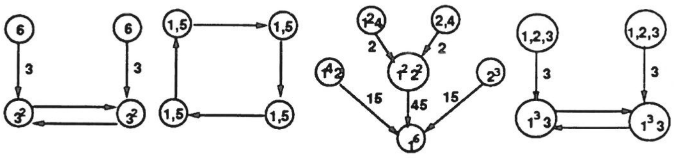

Figure 4. $\mathfrak{S}_6$

Using the above formula, we see that the first shape repeats itself 20 times, the second one 36 times, the third one only once (the component that contains the identity), the fourth one 20 times. The cardinality of the max-f-free subsets in each of the four shapes of components is respectively 6, 2, 210, 6 which gives us that the cardinality $\psi(\mathfrak{S}_6, \sigma \mapsto \sigma^2)$ of the max-f-free subsets of $\mathfrak{S}_6$ is $6 * 20 + 2 * 36 + 210 + 6 * 20 = 522$. Let us conclude by listing the values of $\psi(\mathfrak{S}_n)$ for $1 \leq n \leq 8$: 0, 1, 4, 16, 72, 522, 3 642, 30 753.

References

1. Bruce W. Char, Keith O. Geddes, Gaston H. Gonnet, Benton Leong, Michael B. Monagan, and Stephen M. Watt, *Maple V Library Reference Manual*, Springer-Verlag, New York, Berlin, Heidelberg, 1991.
2. ———, *Maple V Language Reference Manual*, Springer-Verlag, New York, Berlin, Heidelberg, 1991.
3. ———, *First leaves: A tutorial introduction to Maple*, Springer-Verlag, New York, Berlin, Heidelberg, 1991.
4. Y. Fong and Yeong-Nan Yeh, *Near-rings generated by infra-endomorphisms of groups*, To appear.
5. L. K. Hua, *Introduction to number theory*, Springer-Verlag, New York, 1982.
6. Ivan Morton Niven and Herbert S. Zuckerman, *Introduction to the theory of numbers*, John Wiley & Sons, New York, Chichester, Brisbane, Toronto, 1980.
7. W. D. Wallis, Anne Penfold Street, and Jennifer Seberry Wallis, *Combinatorics: Room-squares, sum-free sets, hadamard matrices*, Lecture Notes in Mathematics, vol. 292, Springer-Verlag, New York, 1972.
8. E. Wang, *On double-free sets of integers*, Ars Combinatoria **28** (1989), 97–100.
9. Yeong-Nan Yeh, *On q-free subsets of an operation set*, Submitted.

Pierre Bouchard, Département de mathématiques et d'informatique, Université du Québec à Montréal (UQAM), C.P. 8888, Succursale A, Montréal (Québec), Canada H3C 3P8

Yeong-Nan Yeh, Institute of Mathematics, Academia Sinica, Nankang, Taipei, Taiwan, R.O.C., and Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, U.S.A.