Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org

pp. **X–XX**

DECODING THE MATHIEU GROUP M_{12}

Robert F. Bailey

School of Mathematics and Statistics,
Carleton University, 1125 Colonel By Drive,
Ottawa, Ontario K1S 5B6, Canada

John N. Bray

School of Mathematical Sciences,
Queen Mary, University of London,
Mile End Road, London, E1 4NS, UK

(Communicated by the associate editor name)

Abstract. The sporadic Mathieu group M_{12} can be viewed as an error-correcting code, where the codewords are the group's elements written as permutations in list form, and with the usual Hamming distance. We investigate the properties of this group as a code, in particular determining completely the probabilities of successful and ambiguous decoding of words with more than 3 errors (which is the number that can be guaranteed to be corrected).

1. **Introduction.** In this paper we are concerned with the use of permutation groups as error-correcting codes, with permutations written in list format as the codewords. The use of permutation groups as codes in this way goes back to a 1974 paper of Blake [3], where the use of sharply k-transitive groups was first suggested. (A group G acts sharply k-transitively on a set Ω if for any two [ordered] k-tuples of distinct elements of Ω there is a unique element of G mapping the first to the second.) The idea is developed further in the first author's papers [1, 2], where a decoding algorithm is described.

In particular, we consider the sporadic Mathieu group M_{12}, which acts sharply 5-transitively on 12 points. From a coding theory perspective, this group is of particular interest, since (as Blake indicates in [3]) it produces a code that is roughly comparable to Reed–Solomon codes over $\mathbb{F}_{11}$ and $\mathbb{F}_{13}$ in terms of the length, number of codewords and minimum distance.

Now M_{12} has minimum distance 8, so is guaranteed to correct at most 3 errors. Our main result is to determine completely the probabilities of successful and ambiguous decoding of words with more than 3 errors. This is also of interest, as in practical applications, 100% success is not always required; for instance 90% or 95% may suffice.

2000 *Mathematics Subject Classification.* Primary: 94B60; Secondary: 20D08, 94B25.

Key words and phrases. Permutation code; Mathieu group M_{12}.

The first author was supported in part by an EPSRC CASE studentship sponsored by the UK Government Communications Headquarters (GCHQ).

2. **Definitions and notation.**

2.1. **The Hamming space.** The *Hamming space* $\mathrm{H}(m,n)$ is the set of all ordered m-tuples over the alphabet $\{1,\ldots,n\}$. This is a metric space under the *Hamming distance* d, where $\mathrm{d}(x,y)$ is the number of positions in which x and y differ, for $x,y \in \mathrm{H}(m,n)$. We can make $\mathrm{H}(m,n)$ into a graph by joining $x,y \in \mathrm{H}(m,n)$ just when $\mathrm{d}(x,y)=1$. If X and Y are non-empty subsets of $\mathrm{H}(m,n)$ then $\mathrm{d}(X,Y)$ is defined to be $\min\{\,\mathrm{d}(x,y) : x \in X, y \in Y\,\}$; the *least distance* from $x \in \mathrm{H}(m,n)$ to $\varnothing \neq Y \subseteq \mathrm{H}(m,n)$ is denoted $\mathrm{d}(x,Y)=\mathrm{d}(Y,x)$, and defined to be $\mathrm{d}(\{x\},Y) = \min\{\,\mathrm{d}(x,y) : y \in Y\,\}$. The *minimum distance* of $X \subseteq \mathrm{H}(m,n)$ is defined to be $\min\{\,\mathrm{d}(x,y) : x,y \in X \mid x \neq y\,\}$, with the convention that this be ∞ if $|X| \leqslant 1$.

Elements of $\mathrm{H}(m,n)$ may be regarded as functions from $\{1,\ldots,m\}$ to $\{1,\ldots,n\}$: $(x_1,\ldots,x_m) \in \mathrm{H}(m,n)$ corresponds to the function $f : \{1,\ldots,m\} \to \{1,\ldots,n\}$ that maps i to x_i for all i. If $m=n$ then $\mathrm{H}_n := \mathrm{H}(n,n)$ is the set of functions from $\{1,\ldots,n\}$ to itself, and function composition on H_n makes H_n into a monoid; this is the full transformation monoid of $\{1,\ldots,n\}$. Any permutation group G acting on $\{1,\ldots,n\}$ is a submonoid of H_n (with respect to function composition). In particular, H_n contains the full symmetric group S_n as a submonoid.

Note that pre-multiplication or post-multiplication by elements of S_n preserves Hamming distance: thus $\mathrm{d}(\pi x,\pi y)=\mathrm{d}(x\pi,y\pi)=\mathrm{d}(x,y)$ for all $x,y \in \mathrm{H}_n$, $\pi \in \mathrm{S}_n$. Note also that it is our convention that all maps in H_n shall act on the *right*. Elements of S_n shall be referred to as *permutations*, and elements of H_n are *(Hamming) words*. for all $x \in \mathrm{H}_n$, $\pi \in \mathrm{S}_n$. The full group of distance preserving permutations of H_n is $\mathrm{S}_n \wr \mathrm{S}_n$ of order $(n!)^{n+1}$, whose action on H_n is the so-called product action (see [4, pp. 102–103], for example); the group generated the left and right multiplications is a subgroup $\mathrm{S}_n \times \mathrm{S}_n$ of this. The full automorphism group of $\mathrm{H}(m,n)$ is $\mathrm{S}_n \wr \mathrm{S}_m$ of order $(n!)^m.m!$, again with product action. The subgroup generated by the pre- and post-multiplications is $\mathrm{S}_m \times \mathrm{S}_n$, where the pre- and post-multiplications commute with each other and the pre-multiplications generate a group S_m.

We note that the minimum distance of a non-trivial subgroup $G \leqslant \mathrm{S}_n$ is the least distance from g to $G \setminus \{g\}$, and that this distance is independent of g. Thus the minimum distance of G is $n - \max\{\,\#\mathrm{Fix}(g) : g \in G \mid g \neq \iota\,\}$, this quantity being the minimum number of points moved by a non-identity element of G. Here, and throughout this paper, we use ι to denote the identity permutation.

2.2. **Coding theory terminology.** From a coding theory perspective, the transmitted codeword is a permutation $g \in G \leqslant \mathrm{S}_n$, and the word received is a Hamming word $w \in \mathrm{H}_n$, where the distance $i=\mathrm{d}(g,w)$ is the number of *errors* in w. A word containing i errors can therefore be successfully decoded if there is a unique element of G at distance i from w, and none at distance less than i.

A word w at distance i [with $0 \leqslant i \leqslant n$] from g will *decode correctly* if $\mathrm{d}(w,G) = \mathrm{d}(w,g)=i$, and the only element $h \in G$ such that $\mathrm{d}(w,h)=i$ is $h=g$. The probability that a word at distance i from g will decode correctly is:

$$\mathrm{P}(G,g,i) := \frac{\text{number of words at distance } i \text{ from } g \text{ which decode correctly}}{\text{number of words at distance } i \text{ from } g}.$$

The number of words at distance i from g is $(n-1)^i\binom{n}{i}$, independent of g. Now w decodes g correctly if and only if wg^{-1} decodes $\iota = gg^{-1}$ correctly. Thus the number of words which decode correctly is independent of g, and thus the probability of correct decoding is independent of g. Thus we shall write $\mathrm{P}(G,i)$ instead of

$\mathrm{P}(G,g,i)$. Clearly if G has minimum distance d and $i \leqslant \lfloor\frac{d-1}{2}\rfloor$ then all words at distance i from g will decode correctly, that is $\mathrm{P}(G,i)=1$ for such i.

A word w at distance i from g will *decode incorrectly* if $\mathrm{d}(w,G) < \mathrm{d}(w,g) = i$, and it will *decode ambiguously* if $\mathrm{d}(w,G) = \mathrm{d}(w,g) = i$ and there exists $h \in G \setminus \{g\}$ such that $\mathrm{d}(h,w) = i$. Analogously, we define $\mathrm{Q}(G,i)$ to be the probability that a distance i word decodes ambiguously, and $\mathrm{R}(G,i)$ to be the probability that a distance i word decodes incorrectly. We have $\mathrm{P}(G,i)+\mathrm{Q}(G,i)+\mathrm{R}(G,i)=1$. A word is *green* if it decodes the identity correctly, *yellow* if it decodes the identity ambiguously, and *red* if it decodes the identity incorrectly.

2.3. **The Mathieu group M_{12}.** A *Steiner system* $\mathrm{S}(5,6,12)$ is a set $\mathscr{B}$ of [necessarily 132] 6-element subsets of $\Omega = \{1,\dots,12\}$, such that each 5-element subset of Ω is a subset of precisely one element of $\mathscr{B}$. There is a unique $\mathrm{S}(5,6,12)$ up to permutations of Ω (and there are 5040 altogether). The elements of $\mathscr{B}$ are referred to as *special hexads*, or *hexads* for short. An element $\pi \in \mathrm{S}_{12}$ is an *automorphism* of the $\mathrm{S}(5,6,12)$ $\mathscr{B}$ if $A.\pi = A^\pi \in \mathscr{B}$ for all $A \in \mathscr{B}$. The full automorphism group of any particular $\mathrm{S}(5,6,12)$ is the *Mathieu group* M_{12}, which has size $95040 = 12.11.10.9.8$ and acts sharply 5-transitively on Ω. We note that M_{12} is transitive on subsets of size r for $0 \leqslant r \leqslant 12$, except if $r=6$, when there are two orbits, of sizes 132 and 792. The size 132 orbit consists of the hexads of the associated Steiner system $\mathrm{S}(5,6,12)$.

A particular (standard) copy of M_{12} can be taken to be generated by the permutations

$$\begin{array}{rl} & (1,2,3,4,5,6,7,8,9,10,11), \\ & (1,10)(2,5)(3,7)(4,8)(6,9)(11,12) \\ \text{and} & (3,4)(2,10)(5,9)(6,7). \end{array}$$

This standard copy is the default copy of M_{12} in the computer algebra system GAP, but with different generators. (When working with M_{12} by hand it is usual to use this version of M_{12}, but with the symbols $10,11,12$ relabelled $\mathrm{X},0,\infty$ respectively.) Certain of the following lemmas, especially Lemma 4, use this standard copy of M_{12}. Other lemmas, such as Lemma 2 use a relabelling argument, and therefore use an arbitrary copy of M_{12}.

Since M_{12} is sharply 5-transitive, it has minimum distance 8, as no two elements can agree on more than four points. The sharp 5-transitivity also means that the number of elements [of M_{12}] at distance 8 from a particular element $g \in \mathrm{M}_{12}$ is $(8-1)\times\binom{12}{4} = 7 \times 495 = 3465$.

3. **Words at distance 4 or less.** Since the minimum distance of M_{12} is 8, any $w \in \mathrm{H}_{12}$ satisfying $\mathrm{d}(w,\mathrm{M}_{12}) \leqslant 3$ has a unique nearest neighbour in M_{12}. Thus for $i \leqslant 3$ we have $\mathrm{P}(\mathrm{M}_{12},i)=1$ and $\mathrm{Q}(\mathrm{M}_{12},i)=\mathrm{R}(\mathrm{M}_{12},i)=0$. Furthermore, if $w \in \mathrm{H}_{12}$ and $g \in \mathrm{M}_{12}$ satisfy $\mathrm{d}(w,g)=4$, then $\mathrm{d}(w,\mathrm{M}_{12})=4$, but w need not have a unique nearest neighbour in M_{12}, though g is certainly a nearest neighbour to w in M_{12}. In particular $\mathrm{R}(\mathrm{M}_{12},4)=0$. We now investigate those words w satisfying $\mathrm{d}(w,\mathrm{M}_{12})=4$, especially those not having a unique nearest neighbour in M_{12}.

Lemma 1. *For $g,h \in \mathrm{M}_{12}$ satisfying $\mathrm{d}(g,h)=8$, there are exactly $70=\binom{8}{4}$ words $w \in \mathrm{H}_{12}$ such that $\mathrm{d}(g,w)=\mathrm{d}(h,w)=4$.*

Proof. If $\mathrm{d}(g,h)=8$ then g and h agree in exactly 4 positions, while we require g and w (and h and w) to agree in 8 positions. Thus w agrees with g in 4 of the positions on which g and h differ, and then w agrees with h in the remaining 4 positions on which g and h differ. So w is completely determined by specifying the 4 positions on which g and w agree but g and h differ, and there are $\binom{8}{4}$ such possibilities, each giving rise to a valid w. $\square$

Lemma 2. *Let w be an element of H_{12}. Then there are at most three elements of M_{12} at distance 4 from w.*

Proof. Let g and h be two distinct elements of M_{12} such that $\mathrm{d}(g,w) = \mathrm{d}(h,w) = 4$. Then $\mathrm{d}(g,h) = 8$ and without loss of generality and relabelling points, we may assume that g and h agree on positions 1, 2, 3 and 4, and thus w also agrees on those four positions. Moreover, g and w additionally agree on four more positions, and after relabelling these positions can be taken to be 5, 6, 7 and 8. Now h and w also agree on eight positions, and since g and h agree on precisely four positions, h and w agree on 9, 10, 11 and 12.

Let $k \in \mathrm{M}_{12}$ be distinct from both g and h, and have distance 4 from w. Then k and w agree on eight positions, whereas g, k and w agree on just four, as do h, k and w. The only 8-element subset of $\{1,\ldots,12\}$ intersecting each of $\{1,2,3,4,5,6,7,8\}$ and $\{1,2,3,4,9,10,11,12\}$ in precisely four points is $\{5,6,7,8,9,10,11,12\}$, and thus this is the set on which k and w agree. Since k has been determined on an eight-element subset and no two elements agree more than four points, k is uniquely determined (if it exists at all). $\square$

The above proof shows the following.

Corollary 3. *Let $g,h,k \in \mathrm{M}_{12}$ be distinct elements all at distance 4 from a Hamming word w. The sets on which g & h, g & k and h & k agree are mutually disjoint sets of size 4 partitioning $\{1,\ldots,12\}$.*

We now count the number of ordered tuples (g,h,k,w) such that $g,h,k \in \mathrm{M}_{12}$ with g,h,k distinct, $w \in \mathrm{H}_{12}$ and $\mathrm{d}(g,w) = \mathrm{d}(h,w) = \mathrm{d}(k,w) = 4$. This forces g, h and k to be mutually at distance 8. A couple of reductions are possible. We can postmultiply g,h,k,w by g^{-1}, and thus assume that $g = \iota$, so that now h and k fix 4 points. We can then use the 4-transitivity of M_{12} to conjugate g,h,k,w by a suitable $\pi \in \mathrm{M}_{12}$ so that h fixes 1, 2, 3 and 4. The total number of configurations will be $95040 \times \binom{12}{4}$ multiplied by the number of configurations we do count.

Lemma 4. *The number of ordered tuples (g,h,k,w) where $w \in \mathrm{H}_{12}$; g,h,k are distinct elements of M_{12} and $\mathrm{d}(g,w) = \mathrm{d}(h,w) = \mathrm{d}(k,w) = 4$ such that $g = \iota$ and h fixes 1, 2, 3 and 4 is 18. For all such configurations $w \in \mathrm{S}_{12}$.*

Proof. Throughout, we use the standard copy of M_{12} that we gave in Section 2.3. Calculations such as point stabilisers and membership are easily performed in a computer algebra package such as MAGMA [5] or GAP [7].

By Corollary 3, k fixes a,b,c,d while h and k agree on $\alpha,\beta,\gamma,\delta$, where we have $\{a,b,c,d,\alpha,\beta,\gamma,\delta\} = \{5,6,7,8,9,10,11,12\}$. Now let $\alpha' = \alpha^h = \alpha^k$, $\beta' = \beta^h = \beta^k$, $\gamma' = \gamma^h = \gamma^k$ and $\delta' = \delta^h = \delta^k$. Thus the known values of the functions g,h,k,w are given below.

i	1	2	3	4	a	b	c	d	α	β	γ	δ
$i.g = i^g$	1	2	3	4	a	b	c	d	α	β	γ	δ
$i.h = i^h$	1	2	3	4					α'	β'	γ'	δ'
$i.k = i^k$					a	b	c	d	α'	β'	γ'	δ'
$i.w = i^w$	1	2	3	4	a	b	c	d	α'	β'	γ'	δ'

Now h is a permutation, so $\{1,2,3,4\} \cap \{\alpha', \beta', \gamma', \delta'\} = \varnothing$, and k is also a permutation, so $\{a,b,c,d\} \cap \{\alpha', \beta', \gamma', \delta'\} = \varnothing$. Therefore $\{\alpha', \beta', \gamma', \delta'\} = \{\alpha, \beta, \gamma, \delta\}$, and hence w is a permutation. Thus h stabilises setwise (and thus permutes) $\{a,b,c,d\}$ and k stabilises $\{1,2,3,4\}$.

In our standard copy of M_{12}, the pointwise stabiliser of $\{1,2,3,4\}$, Q say, is a copy of the quaternion group Q_8, consisting of the eight permutations:

$$\begin{array}{cc} \iota & (5,7)(6,11)(8,9)(10,12) \\ (5,6,7,11)(8,10,9,12) & (5,11,7,6)(8,12,9,10) \\ (5,8,7,9)(6,12,11,10) & (5,9,7,8)(6,10,11,12) \\ (5,12,7,10)(6,9,11,8) & (5,10,7,12)(6,8,11,9). \end{array}$$

Now h is a non-identity permutation of Q (the above Q_8), and conjugating h by a suitable power of $\pi = (1,4,2)(6,8,12)(9,10,11) \in \mathrm{Stab}_{M_{12}}(\{1,2,3,4\})$ followed, if necessary, by $(5,8,7,9)(6,12,11,10) \in Q$ we may assume that h is $(5,7)(6,11)(8,9)(10,12)$ or $(5,6,7,11)(8,10,9,12)$. If $h = (5,7)(6,11)(8,9)(10,12)$ then h stabilises just six 4-element subsets of $\{5,6,7,8,9,10,11,12\}$; these are conjugate under $\langle Q, \pi \rangle$ (of order 24), a subgroup which also centralises h. If $h = (5,6,7,11)(8,10,9,12)$ then h stabilises just two 4-element subsets of $\{5,\ldots,12\}$, and these are swapped by $(1,2)(3,4)(6,11)(8,9) \in M_{12}$, an element that also centralises h.

Therefore, in both cases for h we may assume that $\{a,b,c,d\} = \{5,6,7,11\}$. Thus $\{\alpha,\beta,\gamma,\delta\} = \{8,9,10,12\}$. For each of the two possible h we work out $\{\alpha', \beta', \gamma', \delta'\}$ and thus determine the action of k on the eight points $a,b,c,d,\alpha,\beta,\gamma,\delta$, and so determine $k \in M_{12}$ uniquely, if it should exist all. In fact in both cases such a k does exist, and we get k and w as in the table below.

h	$\{a,b,c,d\}$	w	k
$(5,7)(6,11)(8,9)(10,12)$	$\{5,7,6,11\}$	$(8,9)(10,12)$	$(1,4)(2,3)(8,9)(10,12)$
$(5,6,7,11)(8,10,9,12)$	$\{5,7,6,11\}$	$(8,10,9,12)$	$(1,3,4,2)(8,10,9,12)$

The orbits of $(h, \{a,b,c,d\})$ under the conjugations performed have sizes 1×6 (when h has order 2) and 6×2 (when h has order 4). Therefore, we obtain 18 tuples (g,h,k,w) satisfying the conditions of the lemma. □

We now have the tools necessary to prove our main theorem.

Theorem 5. *The probability that a word containing* 4 *errors is decoded uniquely is* $\mathrm{P}(M_{12}, 4) = \frac{14160}{14641} \approx 0.967147.$

Proof. Using the lemmas above, we count possible configurations as follows, where g, h and k are distinct.

- The number of configurations (g,w) where $g \in M_{12}$, $w \in H_{12}$ and $\mathrm{d}(g,w) = 4$ is $|M_{12}| \times \binom{12}{4} \times 11^4 = 95040 \times 495 \times 14641$.
- The number of configurations (g,h,w) where $g,h \in M_{12}$, $w \in H_{12}$ and $\mathrm{d}(g,w) = \mathrm{d}(h,w) = 4$ is $|M_{12}| \times 7\binom{12}{4} \times 70 = 95040 \times 495 \times 490$.
- The number of configurations (g,h,k,w) where $g,h,k \in M_{12}$, $w \in H_{12}$ and $\mathrm{d}(g,w) = \mathrm{d}(h,w) = \mathrm{d}(k,w) = 4$ is $|M_{12}| \times \binom{12}{4} \times 18 = 95040 \times 495 \times 18$.

Multiplying by g^{-1}, we see that the numbers of each of the above configurations in which $g = \iota$ are $\binom{12}{4} \times 14641$, $\binom{12}{4} \times 490$ and $\binom{12}{4} \times 18$ respectively. For $i \in \{1,2,3\}$ let n_i be the number of $w \in H_{12}$ such that $\mathrm{d}(\iota, w) = 4$ and w has exactly i nearest neighbours in M_{12} (by Lemma 2 no such w has 4 or more nearest neighbours in M_{12}). We have:

$$\begin{aligned} n_1 + n_2 + n_3 &= \tbinom{12}{4} \times 14641 \\ n_2 + 2n_3 &= \tbinom{12}{4} \times 490 \\ 2n_3 &= \tbinom{12}{4} \times 18. \end{aligned}$$

Solving these equations gives $n_1 = 14160\binom{12}{4}$, $n_2 = 472\binom{12}{4}$ and $n_3 = 9\binom{12}{4}$. Thus the fraction of w at distance 4 from the identity which do not decode uniquely is

$$\frac{n_2 + n_3}{n_1 + n_2 + n_3} = \frac{481}{14641} \approx 0.032853.$$

Thus the probability $\mathrm{P}(\mathrm{M}_{12}, 4)$ that a given element recieved with 4 errors is guaranteed to decode correctly is approximately 0.967147, and $\mathrm{Q}(\mathrm{M}_{12}, 4) \approx 0.032853$. $\square$

A subtly different question concerns the case when an unknown element of M_{12} is transmitted and received with exactly 4 errors. For $i \in \{1, 2, 3\}$ we let m_i be the number of $w \in \mathrm{H}_{12}$ such that $\mathrm{d}(w, \mathrm{M}_{12}) = 4$. The equations we must now solve are:

$$\begin{aligned} m_1 + 2m_2 + 3m_3 &= |\mathrm{M}_{12}| \times \tbinom{12}{4} \times 14641 \\ 2m_2 + 6m_3 &= |\mathrm{M}_{12}| \times \tbinom{12}{4} \times 490 \\ 6m_3 &= |\mathrm{M}_{12}| \times \tbinom{12}{4} \times 18. \end{aligned}$$

These equations yield $(m_1, m_2, m_3) = (14160C, 236C, 3C)$, where $C = |\mathrm{M}_{12}| \times \binom{12}{4}$. The proportion of such w that are not uniquely decodable is:

$$\frac{m_2 + m_3}{m_1 + m_2 + m_3} = \frac{239}{14399} \approx 0.016598.$$

4. **Words at larger distances.** We now consider the case when we receive a word that has accrued $i \geqslant 5$ errors, with a view to determining how many such words decode correctly, ambiguously or incorrectly. Thus for all $i \geqslant 5$ we wish to determine the values of $\mathrm{P}(\mathrm{M}_{12}, i)$, $\mathrm{Q}(\mathrm{M}_{12}, i)$ and $\mathrm{R}(\mathrm{M}_{12}, i)$. Some of these values for $i = 5, 6, 7$ were done by computer searches using GAP and MAGMA, and the other values were calculated by hand.

Our search will consider those words w having distance i from the identity. Furthermore, since M_{12} acts transitively on i-sets if $i \neq 6$, we shall assume that the positions in which w differs from ι are $\{1, \ldots, i\}$. For the case $i = 6$, we must consider the cases when the positions on which w and ι differ is either $\{1, 2, 3, 4, 5, 6\}$ (a non-hexad in the standard $\mathrm{S}(5, 6, 12)$) or $\{1, 2, 3, 4, 5, 7\}$ (a special hexad). We must also remember to take into account the fact that there are six times as many non-hexads as hexads.

For each of the 11^i possible w, we use the decoding algorithm given in [1] to find all its nearest neighbours in M_{12}, then determine whether w is green, yellow or red. The raw results of our computer search for $i = 5, 6, 7$ are given below. (Case 6H is when the error positions form a hexad, and Case 6N is when they do not.)

Case	#Green	#Yellow	#Red	Total
5	30202	118074	12775	$11^5 = 161051$
6H	132	337740	1433689	$11^6 = 1771561$
6N	66	348571	1422924	$11^6 = 1771561$
7	0	79286	19407885	$11^7 = 19487171$

These translate into (approximate) probabilities as shown below, where we have also given separate conditional probabilities in the case $i = 6$ for when the error positions do or do not form a hexad.

Case i	$P(M_{12}, i)$	$Q(M_{12}, i)$	$R(M_{12}, i)$
5	0.187531	0.733147	0.079323
6H	0.000075	0.190645	0.809280
6N	0.000037	0.196759	0.803204
6	0.000043	0.195886	0.804072
7	0	0.004069	0.995931

For words w, we define $|w|$ to be the number of symbols involved in w; thus $1 \leqslant |w| \leqslant 12$ for $w \in H_{12}$, and $d(w, S_{12}) = 12 - |w|$. If $|w| \geqslant 5$, say w has different symbols in positions $i_1 < i_2 < i_3 < i_4 < i_5$, then the 5-transitivity of M_{12} implies that there is $g \in M_{12}$ which matches w in those positions, and thus $d(w, M_{12}) \leqslant 7$. Similarly, if $|w| = m \leqslant 5$, then w has different symbols in positions $i_1, \ldots, i_m$, and the m-transitivity of M_{12} forces $d(w, M_{12}) \leqslant 12 - m = d(w, S_{12}) \leqslant d(w, M_{12})$, whence $d(w, M_{12}) = 12 - m$. Moreover, if $m \leqslant 4$ the m-point stabiliser is not trivial, and there is more than one element of M_{12} agreeing with w at positions $i_1, \ldots, i_m$. Therefore there are no green words for M_{12} at distance $i \geqslant 8$, and w is a yellow word for M_{12} at distance $i \geqslant 8$ if and only if $|w| = 12 - i$. For distance i, we are considering words w that terminate in $i + 1, \ldots, 12$ at distance i from the identity, and thus the number of yellow words is simply $(12 - i)^i$, out of a total of 11^i such words. Thus we get the following probabilities for $i \geqslant 8$.

i	$P(M_{12}, i)$	$Q(M_{12}, i)$	$R(M_{12}, i)$
8	0	$\frac{4^8}{11^8} \approx 0.000306$	0.999694
9	0	$\frac{3^9}{11^9} \approx 8.35 \times 10^{-6}$	0.999992
10	0	$\frac{2^{10}}{11^{10}} \approx 3.95 \times 10^{-8}$	1.000000
11	0	$\frac{1}{11^{11}} \approx 3.50 \times 10^{-12}$	1.000000
12	0	0	1

5. **Words at distance 7: reducing the search.** The computer search to determine the number of red, yellow and green words at distance i from ι increases significantly in difficulty with increasing i. Firstly, the number of cases we must consider is 11^i (or 2×11^6 when $i = 6$). Secondly, the amount of time required to deal with each case using the decoding algorithm of [1] depends on the number of blocks in a $(12, 5, i)$-uncovering (or a $(12, 7, i)$-covering design, see [8]). For $i = 1, 2, 3, 4, 5, 6, 7$ the sizes of the uncoverings we used were $2, 5, 11, 24, 59, 176, 792$, and the best known lower bounds (as of 26th September 2007) for the sizes of uncoverings with these parameters are $2, 5, 11, 20, 55, 165, 792$, see [8].

The computation for $i = 7$ took about 3 weeks of CPU time on $\approx 3\,\mathrm{GHz}$ computers, a situation we found somewhat unsatisfactory. In contrast, each of the two computations for $i = 6$ required about 9 to 10 hours of CPU time. However, we were able to reduce the computation for $i = 7$ to less than 10 minutes by being able to efficiently eliminate vast swathes of the search space that contained only red words. It is possible that similar reductions may be made for the case $i = 6$. However, it is likely that these will be less effective than those for $i = 7$. The authors did not feel that it would be profitable to pursue this. Another approach to decreasing computer time is discussed in Section 5.1.

Lemma 6. *There are no green words at distance* 7 *from* ι. *Thus* $\mathrm{P}(\mathrm{M}_{12}, 7) = 0$.

Proof. If $\mathrm{d}(w, \iota) = 7$ then $|w| \geqslant 5$. There are then two distinct 5-element subsets $\{i_1, i_2, i_3, i_4, i_5\}$ and $\{j_1, j_2, j_3, j_4, j_5\}$ of $\{1, \ldots, 12\}$ such that w has distinct symbols at positions $i_1, \ldots, i_5$ and at positions $j_1, \ldots, j_5$. Let $g, h \in \mathrm{M}_{12}$ agree with w on positions $i_1, \ldots, i_5$ and $j_1, \ldots, j_5$ respectively.

If $g \neq h$ then w is certainly not green, while if $g = h$ then w and g agree on at least 6 positions, and so w is a red word. □

We consider words w that agree with ι just on positions $8, 9, 10, 11, 12$, and wish to find out how many of these are yellow. The $5^7 = 78125$ such w with $|w| = 5$ are all yellow, and we exclude these from our search by requiring that $j^w \notin \{8, 9, 10, 11, 12\}$ for some j with $1 \leqslant j \leqslant 7$. The following illustrates how we trimmed the search space.

Suppose that $1^w = 2$ (this is one of 42 starting assumptions of the form $j^w = k$ with $1 \leqslant j, k \leqslant 7$ and $j \neq k$ that we must consider). It appears that there are $11^6 = 1771561$ such words to consider. However, w and $(1, 2, 3, 7)(4, 6, 5, 12) \in \mathrm{M}_{12}$ agree on positions $1, 8, 9, 10, 11$. Thus w will be a red word if, for example, $2^w = 3$. The following table gives some permutations of M_{12} (in list format), and the positions on which they are guaranteed to agree with w.

permutation												positions
1	2	3	4	5	6	7	8	9	10	11	12	$8, 9, 10, 11, 12$
2	3	7	6	12	5	1	8	9	10	11	4	$1, 8, 9, 10, 11$
2	4	1	3	6	7	11	8	9	10	5	12	$1, 8, 9, 10, 12$
2	5	10	1	4	3	6	8	9	7	11	12	$1, 8, 9, 11, 12$
2	1	5	7	3	9	4	8	6	10	11	12	$1, 8, 10, 11, 12$
2	7	6	8	1	4	5	3	9	10	11	12	$1, 9, 10, 11, 12$

Therefore if $1^w = 2$ and w is yellow then we have:

$$2^w \in \{6, 8, 9, 10, 11, 12\},\ 3^w \in \{2, 4, 8, 9, 11, 12\},\quad 4^w \in \{2, 5, 9, 10, 11, 12\},$$
$$5^w \in \{2, 7, 8, 9, 10, 11\},\quad 6^w \in \{1, 2, 8, 10, 11, 12\},\ 7^w \in \{2, 3, 8, 9, 10, 12\},$$

which reduces the search space for such w to size $6^6 = 46656$. We can then iterate this process by considering in turn each of the six cases $2^w = 6$, 8, 9, 10, 11 or 12 (with $1^w = 2$ in all six cases). Each iteration of this process takes longer than the previous one since at each stage many more permutations are generated that w must avoid being distance $\leqslant 6$ from. We can use the conjugation action of $\mathrm{Stab}_{\mathrm{M}_{12}}(\{1, 2, 3, 4, 5, 6, 7\}) \cong \mathrm{S}_5$ on these words to speed up the computation. In particular, the only starting configurations we need consider are $1^w = 2$ or 6. The revised version of the program completed in about $1\frac{1}{4}$ minutes and used little memory.

We found 1161 yellow words w with $|w| \geqslant 6$ (and w agreeing with ι on just $\{8, 9, 10, 11, 12\}$). Of these there were 1065 with $|w| = 6$, 96 with $|w| = 7$, and none with $|w| \geqslant 8$. The program, and the data it produced, may be accessed at `http://www.maths.qmul.ac.uk/~jnb/Papers/DecM12/`. The results agree with those obtained from the rather lengthy earlier computation.

5.1. **Conjugation by** M_{12}**.** Another avenue that it is profitable to explore is to utilise the conjugation action of M_{12} on the Hamming words, since this does not alter the distance of a word from ι, nor the colour of the word. Some of this conjugation action is used so that we may assume that the support S of the word is $\{1, \ldots, j\}$ for some j or $\{1, 2, 3, 4, 5, 7\}$. We may then conjugate the words with given support by $H := \mathrm{Stab}_{M_{12}}(S)$, the setwise stabiliser in M_{12} of S. At the suggestion of the referee we pursued this approach. This approach is also required to produce the information (also requested by the referee) in Appendix A and online. We also rewrote the colour determination program to give the extra information needed for the appendix. A side-effect of this is that the program is now somewhat quicker in determining that a word is red (and the above complexity analysis no longer applies); the six error searches now take about 4 hours each and the seven error search takes an estimated 44 hours.[1] However, this approach is not without its problems. What we need to do is to calculate representatives of the H-orbits under conjugation of words having support S, and we prefer that these representatives be canonical in some way. We know of no general algorithm to solve this problem, nor to find the canonical representative of the orbit containing w (for some word w), and our approach was *ad hoc* and unsatisfactory. We made one reduction which speeded up the problem, namely that for such representatives w we may assume that 1^w is a minimal element of an H_1-orbit, where $H_1 = \mathrm{Stab}_H(1)$. In the event, the computations with this method were quick (just over 10 minutes for distance up to 6 and about 65 minutes for distance up to 7), but highly memory intensive: the MAGMA jobs used (at maximum) approximately 90 MB for the distance up to 6 case and 940 MB for the distance up to 7 case. In constrast, the original method and the search space pruning for distance 7 require little memory. Since we recorded the words of each colour (except the red words at distance 7), we are also able to check (and have checked) that these various sets are invariant under the generators of H, thus giving us confidence in our data. This particular check is unavailable with the new method. The programs used here are available online.

6. **Conclusion.** Combining the results of the previous sections, we exhibit the full table of probabilities (all given to 6 decimal places) in Table 1. This is also shown

i	$P(M_{12}, i)$	$Q(M_{12}, i)$	$R(M_{12}, i)$
0	1	0	0
1	1	0	0
2	1	0	0
3	1	0	0
4	0.967147	0.032853	0
5	0.187531	0.733147	0.079323
6	0.000043	0.195886	0.804072
7	0	0.004069	0.995931
8	0	0.000306	0.999694
9	0	0.000008	0.999992
10	0	0.000000	1.000000
11	0	0.000000	1.000000
12	0	0	1

TABLE 1. Probabilities of words of each type

pictorially in Figure 1. From these, we conclude that even though M_{12} is a 3-error correcting code, it is feasible to use it to correct 4 errors with an acceptable

[1]We did about $\frac{1}{11}$ of the latter computation. Namely we considered the 11^6 words w that map 7 to 8 (and fix $8, 9, 10, 11, 12$ and do not fix $1, 2, 3, 4, 5, 6$).

probability of decoding uniquely. For 5 errors, M_{12} is not feasible as an error-correcting code, as the probability $\mathrm{P}(\mathrm{M}_{12}, 5)$ is too small. However, the *detection* of 5 errors is feasible, as although the probability of decoding uniquely is fairly small, the probability $\mathrm{R}(\mathrm{M}_{12}, 5)$ of decoding incorrectly is even smaller. From 6 errors onwards, the use of M_{12} for either detection or correction is not feasible.

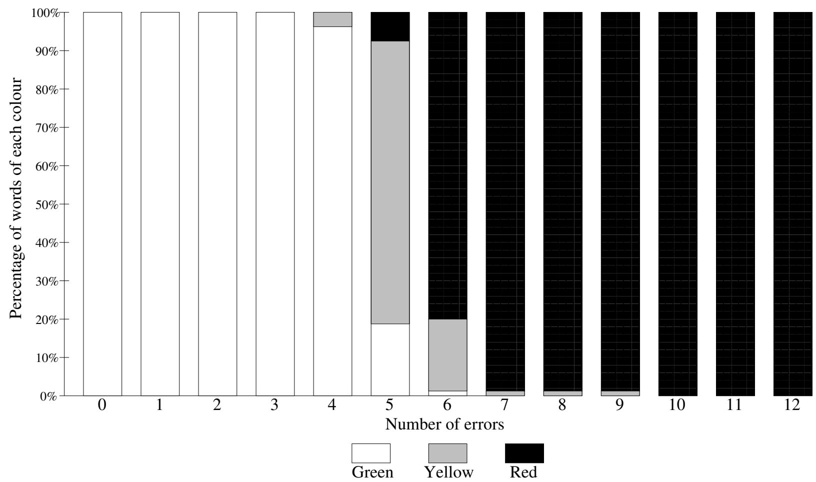

FIGURE 1. Percentage of words of each colour

Appendix A. **Other useful information.** At the suggestion of the referee, we produce some information about words at small distance from ι, and the conjugation action of M_{12} thereon. (Note that conjugation by elements of M_{12} preserves the properties we wish to tabulate.) This information was calculated with the aid of a computer. There is one line in the table for each orbit of words of distance at most 3 from ι, as well as all non-green (necessarily yellow) words at distance 4. The #Nei column gives the number of nearest neighbours to w in M_{12}, and the orbit size column gives numbers in the form $M \times N$ where M is the number of M_{12} images of $\mathrm{supp}(w)$, the support of w, and N is the number of conjugates of w under $\mathrm{Stab}_{\mathrm{M}_{12}}(\mathrm{supp}(w))$. The orbit representatives w given are minimal under the lexicographic order, subject to w having support $\{1, \ldots, i\}$ for some i or support $\{1, 2, 3, 4, 5, 7\}$.

We have calculated these tables for words at distance up to 7 from ι. For the larger distances, these tables become unwieldy. The number of rows these tables have for the larger distances are 13 (all green) for distance 3; 109 (101 green, 8 yellow) for distance 4; 1431 (280 green, 1030 yellow, 121 red) for distance 5; 2607 (3 green, 524 yellow, 2080 red) for distance 6, hexad support; 15059 (3 green, 3026 yellow, 12030 red) for distance 6, non-hexad support; and 163251 (712 yellow, 162539 red) for distance 7. These extended tables are available electronically at `http://www.maths.qmul.ac.uk/~jnb/Papers/DecM12/`.

REFERENCES

[1] R. F. Bailey, *Error-correcting codes from permutation groups*, Discrete Math., submitted.
[2] R. F. Bailey, *Uncoverings-by-bases for base-transitive permutation groups*, Des. Codes Cryptogr., **41** (2006), 153–176.

$d(w,\iota)$	Orbit size	Colour	$d(w, M_{12})$	#Nei	Representative w
0	1×1	Green	0	1	1 2 3 4 5 6 7 8 9 10 11 12
1	12×11	Green	1	1	2 2 3 4 5 6 7 8 9 10 11 12
2	66×1	Green	2	1	2 1 3 4 5 6 7 8 9 10 11 12
2	66×20	Green	2	1	2 3 3 4 5 6 7 8 9 10 11 12
2	66×10	Green	2	1	3 3 3 4 5 6 7 8 9 10 11 12
2	66×90	Green	2	1	3 4 3 4 5 6 7 8 9 10 11 12
3	220×6	Green	3	1	2 1 1 4 5 6 7 8 9 10 11 12
3	220×27	Green	3	1	2 1 4 4 5 6 7 8 9 10 11 12
3	220×2	Green	3	1	2 3 1 4 5 6 7 8 9 10 11 12
3	220×54	Green	3	1	2 3 4 4 5 6 7 8 9 10 11 12
3	220×27	Green	3	1	2 4 2 4 5 6 7 8 9 10 11 12
3	220×54	Green	3	1	2 4 4 4 5 6 7 8 9 10 11 12
3	220×432	Green	3	1	2 4 5 4 5 6 7 8 9 10 11 12
3	220×9	Green	3	1	4 4 4 4 5 6 7 8 9 10 11 12
3	220×216	Green	3	1	4 4 5 4 5 6 7 8 9 10 11 12
3	220×216	Green	3	1	4 5 6 4 5 6 7 8 9 10 11 12
3	220×72	Green	3	1	4 5 7 4 5 6 7 8 9 10 11 12
3	220×144	Green	3	1	4 5 9 4 5 6 7 8 9 10 11 12
3	220×72	Green	3	1	4 5 11 4 5 6 7 8 9 10 11 12
4	495×3	Yellow	4	3	2 1 4 3 5 6 7 8 9 10 11 12
4	495×48	Yellow	4	2	2 1 5 7 5 6 7 8 9 10 11 12
4	495×6	Yellow	4	3	2 3 4 1 5 6 7 8 9 10 11 12
4	495×192	Yellow	4	2	2 3 5 11 5 6 7 8 9 10 11 12
4	495×96	Yellow	4	2	2 5 4 10 5 6 7 8 9 10 11 12
4	495×96	Yellow	4	2	2 5 11 9 5 6 7 8 9 10 11 12
4	495×24	Yellow	4	2	5 7 10 12 5 6 7 8 9 10 11 12
4	495×16	Yellow	4	2	5 9 6 12 5 6 7 8 9 10 11 12

TABLE 2. Extra information on words close to ι

[3] I. F. Blake, *Permutation codes for discrete channels*, IEEE Trans. Inform. Theory, **20** (1974), 138–140.

[4] P. J. Cameron, "Permutation Groups," London Mathematical Society Student Texts **45**, Cambridge University Press, Cambridge, 1999.

[5] J. J. Cannon *et al.*, The MAGMA programming language, Version 2.11, School of Mathematics and Statistics, University of Sydney (2004).

[6] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker and R. A. Wilson, "An ATLAS of Finite Groups," Clarendon Press, Oxford, 1985 (reprinted with corrections 2003).

[7] The GAP Group, GAP—Groups, Algorithms, and Programming, Version 4.4, (2004).

[8] D. M. Gordon, "La Jolla Covering Repository," `http://www.ccrwest.org/cover.html`.

Received June 2007; revised September 2007.

E-mail address: robertb@math.carleton.ca
E-mail address: J.N.Bray@qmul.ac.uk