Recent results on covering problems

RECENT RESULTS ON COVERING PROBLEMS

J.H. van Lint*

1. Introduction

The aim of this paper is to do three things:

(i) To list the papers on covering radius problems that have come to my attention since the appearance of the survey paper by Cohen et al. in 1985 (cf. [10]) and to give some idea of the type of problems treated in these papers. We point out that many of these papers have not appeared yet;

(ii) To discuss a few of the most interesting new ideas occurring in the papers mentioned above;

(iii) To survey the recent ($\geq$ 1983) developments on the so-called football pool problem.

We shall assume that the reader is familiar with coding theory and the covering radius problem (see the survey [10]). We shall use the following terminology. We consider an alphabet Z (in this paper the alphabet will nearly always be one of the fields $\mathbb{F}_2$ or $\mathbb{F}_3$).

A *code* C of length n is a subset of Z^n. We use the following notation. A k-dimensional linear subspace of $\mathbb{F}_q^n$ is called an $[n, k]$ code. If the minimum distance of the code C is d, then we shall call C an $[n, k, d]$ code. If C is not linear and $|C| = K$, then we use (n, K) code resp. (n, K, d) code. If $\mathbf{x} \in Z$ then we define the *sphere* of radius R around $\mathbf{x}$ by

$$B_R(x) := \{y \in Z^n : d(\mathbf{x}, \mathbf{y}) \leq R\}. \tag{1.1}$$

If $R = 1$ such a sphere is often called a *rook-domain* (because the case $k = 8, n = 2, R = 1$ corresponds to the positions a rook on a chessboard can reach in at most one move).
Let $|Z| = q$. Then we have

$$V_q(n, R) := |B_R(\mathbf{x})| = \sum_{i=0}^{R} \binom{n}{i} (q-1)^i. \tag{1.2}$$

We shall say that C is an *R-covering* of Z^n if every word in Z^n has distance at most R to some codeword, i.e.

Department of Mathematics and Computing Science, Eindhoven University of Technology, Eindhoven, Netherlands.
This research was supported in part by the Institute of Mathematics and its Applications with funds provided by the National Science Foundation.

(1.3) $$Z^N = \bigcup_{c \in C} B_R(c).$$

The minimal value of R for which this is true is called the *covering radius* of C. For a linear code the covering radius is equal to the weight of a coset leader of maximal weight.

Another very useful definition is given by using the parity check matrix of the code. If H is the $n-k$ by n parity check matrix of an $[n, k]$ code C, then the covering radius R of C is the smallest integer t such that every syndrome is a linear combination of at most t columns of H. We point out that some authors use $t(C)$ for the covering radius of C, some others use $C\ R(C)$, and furthermore that the notation $(n, K)\ R$ code is used for an (n, K) code with covering radius R.

2. Functions related to the covering radius.

Several functions have been introduced to express information about the covering radius of codes. The best known are

(2.1) $t[n, k] :=$ minimal covering radius for $[n, k]$ codes.

(2.2) $t(n, K) :=$ minimal covering radius for (n, K) codes.

(2.3) $K_q(n, R) := \min\{|C| : C \text{ is an } R\text{-covering of } Z^n\}$.

(We usually omit the index q if it is 2).

(2.4) $k(n, R) :=$ minimal dimension of a linear code of length n and covering radius R.

The following obvious inequality is known as the *sphere covering bound*:

(2.5) $$K_q(n, R) \geq \frac{q^n}{V_q(n, R)}.$$

In a recent paper ([4]) Brualdi, Pless and Wilson introduced a new function, which they called the *length function*, as follows:

(2.6) $l(m, r) :=$ smallest length of a binary code of codimension m and covering radius r.

At first, this does not seem too useful because one can simply translate most of the known results for $t[n, k]$ to obtain the equivalent statements for $l(m, r)$. However, in tables the function $l(m, r)$ saves a lot of space. E.g. the information $t[k+8, 8] = 2$ for $k = 18, 19, \cdots, 246$ takes many entries 2 in a table, where $l(8, 1) = 255$, $l(8, 2) \leq 26$ tells us the same thing. The translation of (2.5) is

(2.7) $$l(m, r) \geq \min\left\{n : \sum_{i=0}^{r} \binom{n}{i} \geq 2^m\right\}$$

and the expression on the right hand side is called the *first feasible length* for a code of codimension m and covering radius r.

3. Normal codes and the ADS construction.

Two recent long papers on covering radius with many new results and methods are [14] by Graham and Sloane and [11] by Cohen, Lobstein and Sloane. From these we quote the following.

DEFINITION 3.1. Let C be a binary $[n, k]R$ code. For $1 \le i \le n$ denote by $C_0^{(i)}$ resp. $C_1^{(i)}$ the subcode of C consisting of the codewords with i-th coordinate 0 resp. 1. We assume that both are nonempty, i.e. they each contain half of the codewords. Define

$$N^{(i)} := \max \{d(\mathbf{x}, C_0^{(i)}) + d(\mathbf{x}, C_1^{(i)}) : \mathbf{x} \in \mathbb{F}_2^n\} \tag{3.1}$$

and call this the norm of C with respect to position i. Then define

$$N := \min \{N^{(i)} : 1 \le i \le n\}. \tag{3.2}$$

We call N the norm of the code and coordinate positions i for which $N = N^{(i)}$ are called *acceptable*. The code C is called a *normal* code if

$$N \le 2R + 1. \tag{3.3}$$

DEFINITION 3.2. For a code C over a q-ary alphabet we define "normal" in a similar way. In (3.1) we now have q subcodes $C_j^{(i)}$ depending on the value of the i-th coordinate and (3.3) must be replaced by

$$N \le qR + (q-1). \tag{3.4}$$

So, a code is normal if for some i and every word $\mathbf{x}$ the average distance from $\mathbf{x}$ to the subcodes $C_j^{(i)}$, $0 \le j \le q-1$, is less than $R+1$.

Whereas most binary codes seem to be normal (cf. [14]) it is quite difficult to find nonbinary normal codes!

EXAMPLE 3.1. Let C be a perfect q-ary Hamming code ($R = 1$). If we take $\mathbf{x} = 0$ in the q-ary analog of (3.1), then we find

$$N = 0 + (q-1) \times 3 = 3q - 3,$$

whereas $qR + (q-1) = 2q - 1$, which is less unless $q = 2$. (In fact $N^{(i)} = 3q - 3$ for all $\mathbf{x}$.)

DEFINITION 3.3. Let C_i be an $[n_i, k_i]R_i$ code ($i = 1, 2$). We assume that the last coordinate of C_1 is acceptable and that the first coordinate of C_2 is acceptable. The *amalgamated direct sum* $C_1 \dot{+} C_2$ of C_1 and C_2 is the $[n_1+n_2-1, k_1+k_2-1]$ code consisting of the words $(c_1, c_2, \cdots, c_{n_1+n_2-1})$ for which the word $(c_1, c_2, \cdots, c_{n_1})$ is in the code C_1 and $(c_{n_1}, c_{n_1+1}, \cdots, c_{n_1+n_2-1})$ is in C_2.

This construction is called the "ADS-construction".

THEOREM 3.1. *If C_1 and C_2 are normal then the covering radius of $C_1 \dot{+} C_2$ is $R_1 + R_2$.*

In [14] this is called "saving a coordinate" (compared to the usual direct sum construction). Methods are given that save more than one coordinate.

The paper [11] is devoted to bounds for $K(n, R)$ for binary codes. Some of the results are based on the (obvious) generalization of Definition 3.1 to nonlinear codes. The lower bounds are improvements of (2.5) obtained by estimating how many codewords are "covered" more than once by the code, i.e. have distance $\leq R$ to more than one codeword. Among the interesting results are the following

$$K(5, 1) = 7, \quad K(6, 1) = 12, \quad K(11, 1) \leq 192,$$

and

$$K(2R + 3, R) = 7, \quad K(2R + 4, R) \leq 12.$$

One of the useful ideas in the constructions is the concept of a *piecewise constant code*. This is a code of length $n_1 + n_2 + \cdots + n_t$ for which the coordinates are partitioned into blocks of size n_i, $(1 \leq i \leq t)$, and any permutation of any of the blocks is an automorphism of the code.

EXAMPLE 3.2. $K(5, 1) \leq 7$ is demonstrated by the piecewise constant code

0 0 0 0 0
0 0 1 1 1
1 0 0 0 0
0 1 0 0 0
1 1 0 1 1
1 1 1 0 1
1 1 1 1 0

The following construction is given in [11] and attributed to Katsman and Litsyn, and to Mollard. Actually this idea already occurred in a slightly different form in [22]. (See Example 7.1.)

THEOREM 3.2. *Let C be an (n, K) code with covering radius 1. We define C^* to be the code with codewords $(c_0, c_1, \cdots, c_n, c'_1, \cdots, c'_n)$, for which $c_0 + c_1 + \cdots + c_n = 0$ and $(c'_1 - c_1, \cdots, c'_n - c_n) \in C$. Then C^* is a $(2n + 1, 2^n K)$ code with covering radius 1.*

The idea of the proof is the same as in Example 7.1.

4. New bounds.

It is not surprising that many of the recent papers in this area are on bounds for the covering radius functions. Quite often they concern ad hoc results that we shall only briefly mention. However, some nice new ideas have also been introduced in these papers. In Section 3 we mentioned some improvements of the bound (2.5). The most significant recent improvements to the sphere covering bound are due to van Wee [36]. We do not mention all the results (for these see the tables for $K(n, R)$ in the paper) but illustrate the main idea by one simple example.

THEOREM 4.1. *If n is even then $K(n, 1) \geq 2^n/n$.*

Proof. Let n be even and let C be a code with covering radius 1 in $\mathbb{F}_2^n$, $C' := \mathbb{F}_2^n \setminus C$, and let A be the set of words in $\mathbb{F}_2^n$ that have distance ≤ 1 to at least two words of C. Note that if $d(\mathbf{x}, \mathbf{c}) = 1$ or 2, then $| B_1(\mathbf{x}) \cap B_1(\mathbf{c}) | = 2$ and since $| B_1(\mathbf{x}) | = n + 1$ is odd, it follows that if $\mathbf{x} \in C'$, then $| B_1(\mathbf{x}) \cap A | \geq 1$. Now count in two ways the pairs $(\mathbf{x}, \mathbf{y})$ with $\mathbf{x} \in C'$, $\mathbf{y} \in A$, $d(\mathbf{x}, \mathbf{y}) \leq 1$. As we saw above, this number is at least $| C' | = 2^n - | C |$. On the other hand, this number is trivially at most $| A | \cdot (n-1)$. Therefore $| A | \geq (2^n - | C |) / (n-1)$. Since $| C | (n+1) \geq 2^n + | A |$ the result follows. □

Most of the improved bounds in [36] are based on the generalization of Theorem 4.1 to the case of $R > 1$.

The next result also concerns a significant improvement of earlier bounds. In fact, the following construction due to Wilson [4] gives an exponential improvement on the upper bounds found by methods using generalizations of the ADS-construction. Again we only illustrate the idea by giving one example. We are interested in covering radius $R = 2$. Using the terminology of the length function $l(m, r)$, we wish to find a subset S of $\mathbb{F}_2^m$ of minimal size such that any element of $\mathbb{F}_2^m$ is the sum of at most two elements of S. Considering everything projectively, this can be reformulated as follows. Find a set S of minimal cardinality in the space $\mathbf{P} = PG(m-1, 2)$ such that every point of $\mathbf{P}$ is on a secant of S. In general, a set S in $PG(n, q)$ such that every point of the space is on a secant of S is called a "secant covering set".

LEMMA 4.1. *Let $q = 2^a$. Then there is a secant covering set in $PG(3, q)$ with $2q + 1$ points.*

Proof. Let π be a plane in $PG(3, q)$ and let O be an oval in π. (So $| O | = q + 1$.) We consider a line l through the nucleus N of O but not in π. It is trivial to see that $O \cup l \setminus N$ is a secant covering set with $2q + 1$ points: for points in π use the definition of oval and the fact that N is on l; for a point not on π use the fact that the plane through such a point and l intersects π in a line that must meet O. □

DEFINITION 4.1. Let S be the 4 by $2q + 1$ matrix that has as columns the points of the secant covering set of Lemma 4.1. Let α be a primitive element of $\mathbb{F}_q$ and let M be the matrix

$(1\ \alpha\ \cdots\ \alpha_{q-2})$. We define a *binary* matrix S^* as follows: in the Kronecker product $S \otimes M$ replace each entry (an element of $\mathbb{F}_q$) by its representation as a column vector in $\mathbb{F}_2^a$.

THEOREM 4.2. *The columns of S^* are a secant covering set in $PG(4a-1, 2)$.*

Proof. An arbitrary column vector in $\mathbb{F}_2^{4a}$ can be interpreted as a column vector in $\mathbb{F}_q^4$, which by the definition of S is a linear combination of two columns of S and therefore the sum of two columns of S^*, since every nonzero multiple of a column of S is a column of S^*. □

The following theorem (from [4]) is an immediate consequence.

THEOREM 4.3. *For $a \geq 1$*

$$l(4a, 2) \leq (2^{a+1}+1)(2^a-1). \tag{4.1}$$

Note that an ADS-type construction saving e columns and using two Hamming codes would give the bound $l(4a, 2) \leq 2^{2a+1} - 2 - e$ and the two bounds differ by $2^a - e - 1$.
In [4] it is shown that some of the results of Graham and Sloane [14] can be translated as follows for the length function:

$$l(2s+1, s) = 2s+5 \quad \text{for} \quad s \geq 1, \tag{4.2}$$

$$l(2s, s-1) = 2s+6 \quad \text{for} \quad s \geq 4. \tag{4.3}$$

Since $l(m, 1)$ follows from the Hamming codes, the only difficult values of $l(m, r)$ are in the interval $2 \leq r < \frac{1}{2}m - 1$. The most recent update on these values can be found in a preprint by Brualdi and Pless [6]. The lower bounds

$$l(9, 2) \geq 33, \quad l(10, 2) \geq 46, \quad l(12, 2) \geq 91$$

are from [4], the result

$$l(9, 3) \geq 16$$

due to Simonis [31] was reported at this meeting, for $4 \leq r \leq 8$ there are several bounds due to van Wee as mentioned above, and finally

$$l(7, 2) \geq 18, \quad l(8, 2) \geq 24$$

occur in a recent preprint of Calderbank and Sloane [9]; the second one also occurs in [31]. Besides the upper bounds already mentioned in this section there are a few unpublished results found by computer. These are: $l(8, 2) \leq 26$, $l(9, 2) \leq 41$, $l(9, 3) \leq 18$, and $l(12, 3) \leq 38$. In [6] these are quoted as private communications by D. Ashlock and R. Kibler.

A lower bound on $K(n, R)$ from [11] is improved by van Lint, jr. in [27] as follows. Let C have length $n+2$, covering radius $R+1$, and assume that $|C| = M < K(n, R)$. Then for all

$1 \le i < j \le n+2$, the pair (c_i, c_j), where $\mathbf{c}$ runs through C, takes on all four possible values. Let A_j be the subset of C consisting of codewords with $c_j = 1$. Then by the assertion above, these sets have the property that if $k \neq l$ then all four sets $A_k \cap A_l$, $A_k \cap \bar{A}_l$, $\bar{A}_k \cap A_l$, $\bar{A}_k \cap \bar{A}_l$ are nonempty. A well known result from extremal set theory then states that

$$n+2 \le \left(\begin{matrix} M-1 \\ \lfloor \frac{M}{2} \rfloor - 1 \end{matrix} \right).$$

A consequence of this is the bound

$$K(2R+5) \ge 8 \quad \text{for} \quad R \ge 6.$$

As we saw above, the lower bounds on $K(n, R)$ of [14] and [36] were found by estimating the number of words that are covered more than once. More results of this type are given by Honkala in [17], e.g. $K(8, 2) \ge 10$, $K(11, 3) \ge 11$, $K(14, 4) \ge 13$.

In [18] Honkala and Hämäläinen construct new covering codes (and improve some upper bounds for $K(n, R)$) by first constructing the code of length n consisting of all words with weight w, where $w \in W$, W being a set satisfying a number of special conditions, and then doubling the length by the rules $0 \to 00$, $1 \to 11$ or $0 \to 10$, $1 \to 01$.

The well known *Griessmer bound* for binary codes states that if C is an $[n, k, d]$ code, then

$$n \ge \sum_{i=0}^{k-1} \lceil \frac{d}{2^i} \rceil .$$

The covering radius of codes meeting this bound is considered in [7] by Busschbach, Gerretzen and van Tilborg.

In [19] Janwa proves the following similar inequality for the covering radius. Let C be as above and assume that C has covering radius R. Then

$$R \le n - \sum_{i=1}^{k} \lceil \frac{d}{2^i} \rceil .$$

Let G be a generator matrix for an $[n, k]$ binary code C. Let G have a distinct columns, none of which is zero, and let these columns have multiplicities $m_1, m_2, \cdots, m_a$. In [33] Sloane defines the *normalized covering radius* ρ of C to be

$$\rho = R - \sum_{i=1}^{a} \lfloor \frac{m_i}{2} \rfloor .$$

In this paper and in Kilby and Sloane [23] the covering radius of codes of low dimension is studied and upper bounds are obtained for the normalized covering radius.

5. Subcodes.

We mention four papers on the covering radius of subcodes, three of which are related. In [2] Adams shows that if C_0 is a subcode of codimension 1 of a binary linear code C with covering radius R, then the covering radius R_0 of C_0 satisfies $R_0 \leq 2R+1$. In [5] this is generalized to $R_0 \leq (i+1)R+i$ in case the codimension is i. Calderbank [8] shows that $R_0 \leq 2R+2^i-1$.

In [20] Janwa and Mattson consider even-weight subcodes and their covering radius.

6. Special topics.

For the sake of making this survey as complete as the present author can do at this time, we list a number of papers that we know about but have not seen or have not studied sufficiently.
The references [12], [20], [28] and [34] are about special codes, such as cyclic codes, BCH codes, RM codes, etc.
The references [16] and [27] concern codes with *mixed alphabets*, i.e. some symbols are from $\mathbb{F}_2$, others from $\mathbb{F}_3$.

The first sections of this paper have clearly demonstrated that very much has been going on recently concerning the covering radius of binary codes and also that the present author cannot keep up with the reading of all this interesting material! In the following sections we turn to the problem of covering radius 1 for ternary codes.

7. The football pool problem.

In one entry in a football pool one forecasts the outcome of n football matches (win, lose or draw). To win first prize all forecasts have to be correct. If one wishes to guarantee winning the first prize, no matter what the outcome of the matches is, then one obviously has to submit 3^n entries. To win second prize one of the forecasts may be incorrect. In order to guarantee second prize, we consider each entry as a word in $\mathbf{F}_3^n$. If the set of entries is a 1-covering of the space, then winning at least the second prize again does not depend on the outcome of the matches. The number of entries needed is usually denoted by $\sigma(n, 3)$. So, in the terminology of (2.3) we have

(7.1) $\qquad \sigma(n, 3) = K_3(n, 1).$

Since the ternary Hamming codes $\mathbb{H}_4$, $\mathbb{H}_{13}$ of length 4 resp. 13 (both perfect codes) yield $\sigma(4, 3) = 9$, $\sigma(13, 3) = 3^{10}$, we have

(7.2) $\qquad \sigma(n, 3) \leq 3^{n-2}$ for $5 \leq n \leq 12.$

The research in this area (in the last 25 years!) has been concerned with the values in (7.2). In [21] it was shown that equality holds in (7.2) for $n = 5$; the proof is very long. Of course trivially

(7.3) $\sigma(n+1, 3) \leq 3\sigma(n, 3)$,

so an improvement of the upper bound for one value of n also could lead to an improvement for larger values.

We start by explaining a simple construction (cf. [22]) that is related to the one mentioned in Theorem 3.2. It still holds the world record for $n = 9$.

EXAMPLE 7.1. $\sigma(9, 3) \leq 2 \cdot 3^6$
The proof of this assertion is by construction. The code C is

$$C := \{(x_0, x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4\},$$

where

(i) $\mathbf{x} \in \mathbb{F}_3^4$,

(ii) $\sum_{i=0}^{4} x_i \neq 0$,

(iii) $\mathbf{x} - \mathbf{y} \in \mathbb{H}_4$.

Actually the description of C in [22] was as follows. Each point of $\mathbb{F}_3^5$ was given one of 9 colors or the color "blank" in such a way that each blank point had distance 1 to a point of each of the 9 colors. The points were then replaced by an empty copy of $\mathbb{F}_3^4$ in the case of blank and a copy containing one of the nine cosets of $\mathbb{H}_4$ (one corresponding to each color) otherwise. This description of C is generalized in [3].

It is easy to see (from either description) that C has covering radius 1.

Using combinatorial constructions several new upper bounds for $\sigma(n, 3)$ for $6 \leq n \leq 8$ were found in the last five years. In 1983 Weber [35] proved $\sigma(6, 3) \leq 79$ and Fernandes and Rechtschaffen [13] showed that $\sigma(7, 3) \leq 225$ and $\sigma(8, 3) \leq 567$. These results were obtained using 2-step coverings (like our second description of C in Example 7.1).

In 1984 Blokhuis and Lam [3] generalized the idea of Example 7.1 as follows.

DEFINITION 7.1. Let $A = (I\ M) = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n)$ be a $r \times n$ matrix, where I is the $r \times r$ identity matrix and M has entries from $\mathbb{F}_3$. A subset S of $\mathbb{F}_3^r$ is said to *cover* $\mathbb{F}_3^r$ *using* A if

$$\mathbb{F}_3^r = \{\mathbf{s} + \alpha \mathbf{a}_i : \mathbf{s} \in S, \alpha \in \mathbb{F}_3, 1 \leq i \leq n\}.$$

THEOREM 7.1. *If S is a covering of $\mathbb{F}_3^r$ using A, then $W = \{\mathbf{w} \in \mathbb{F}_3^n : A\,\mathbf{w} \in S\}$ is a 1-covering of $\mathbb{F}_3^n$ with $|W| = |S| 3^{n-r}$.*

Proof. Let $\mathbf{x} \in \mathbb{F}_3^n$. Then $A\,\mathbf{x} = \mathbf{s} + \alpha\,\mathbf{a}_i$. Therefore, if $\mathbf{e}_i$ is the i-th basis vector in $\mathbb{F}_3^n$, $A(\mathbf{x} - \alpha\,\mathbf{e}_i) \in S$, so $\mathbf{x} - \alpha\,\mathbf{e}_i \in W$. □

Using Theorem 7.1 Blokhuis and Lam proved that $\sigma(7, 3) \leq 216$ and by a similar construction $\sigma(10, 3) \leq 5 \cdot 3^6$.
The idea of Definition 7.1 was generalized by van Lint, jr. in [27].

The most recent result is a combinatorial construction showing that $\sigma(8, 3) \leq 6 \cdot 3^4$. Again Theorem 7.1 is used, with $r = 4$, and

$$M = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} M_1 \\ 0\,0\,0\,0 \end{bmatrix}.$$

Two points $\mathbf{x}$ and $\mathbf{y}$ are said to form a "*pair*" if $x_i = y_i$ for one value of i and $x_i = y_i - 1$ for the other two values of i. It is easy to find three such pairs such that each of them covers $\mathbb{F}_3^3$ using $(I\ M_1)$ with the exception of the other two pairs. This then immediately yields a set of six points in $\mathbb{F}_3^4$ covering that space using $(I\ M)$. This result is given in [26] with several other new bounds for $\sigma(n, 3)$, all found using simulated annealing (to be discussed in the next section). In the present case a close analysis of the computer output led to the idea described above.

8. Simulated annealing.

In [26] van Laarhoven, Aarts, van Lint and Wille present three new upper bounds: the bound for $\sigma(8, 3)$ mentioned in the previous section and furthermore $\sigma(6, 3) \leq 73$ and $\sigma(7, 3) \leq 186$. These bounds were obtained using a fairly new technique known as *simulated annealing* or *statistical cooling*. For an extensive treatment of this technique we refer to the book [25] by van Laarhoven and Aarts or the survey paper [1] by the same authors. A good introduction to applications in coding theory is given by van der Ham [15]. The essential difference between simulated annealing and the standard iterative improvement algorithms used to minimize certain functions is easily described in one sentence. In the standard algorithm a transition from one state (or configuration) to some neighbouring state is accepted only if the value of the function to be minimized decreases due to this transition; in simulated annealing transitions that increase this function are sometimes accepted and the decision is governed by a probabilistic algorithm. To give a somewhat more detailed idea of the background we quote from the introduction of [15].

Annealing is a technique from statistical mechanics; solids are annealed by raising the temperature to a maximal value at which the particles randomly arrange in the liquid phase, followed by cooling to force the particles into a low energy state of a regular lattice. At each temperature T, the solid is allowed to reach *thermal equilibrium*, characterized by a probability distribution of being in a state with energy E given by the *Boltzmann distribution*:

$$Pr\{\mathbf{E}=E\} = \frac{1}{Z(T)} \cdot \exp\left(-\frac{E}{k_B T}\right)$$

where $Z(T)$ is a normalization factor and k_B the Boltzmann constant. As the temperature decreases, the Boltzmann distribution concentrates on states with lowest energy and finally, when the temperature approaches zero, only minimum energy states have a non-zero probability of occurrence. However, if the cooling is too rapid, i.e. if the solid is not allowed to reach thermal equilibrium for each temperature value, defects can be "frozen" into the solid and a minimum energy state cannot be reached.

To simulate the evolution to thermal equilibrium of a solid for a fixed value of the temperature T, Metropolis et al. [29] proposed a *Monte Carlo method*, which generates sequences of states of the solid in the following way. Given the current state of the solid, characterized by the positions of its particles, a small, randomly generated perturbation is applied by a small displacement of a randomly chosen particle. If the difference in energy, ΔE, between the slightly perturbed state and the current one is negative, i.e. if the perturbation results in a lower energy for the solid, then the process is continued with the new state (the new state is accepted). If $\Delta E > 0$, then the probability to accept the perturbed state is given by

$$\exp(-\Delta E / k_B T).$$

This acceptance rule for new states is referred to as the *Metropolis criterion*. Following this criterion, the system eventually evolves into thermal equilibrium, i.e. after a large number of perturbations the probability distribution of the states approaches the Boltzmann distribution. This method is known as the *Metropolis Algorithm*.

Kirkpatrick et al. [24] used this Metropolis Algorithm to solve combinatorial optimization problems by using a *cost function* in place of the energy and configurations in place of the states of the solid; the temperature then assumes the role of a control parameter that is no longer fixed on one value.

We illustrate how the method can be used to show that $\sigma(6, 3) \leq 73$. The following algorithm was executed a number of times using a code C with $|C| = \sigma$, starting with $\sigma = 80$ and then decreasing the value of σ until the algorithm failed to find a covering code. This happened for $\sigma = 72$.

The algorithm starts by taking a random code C of the prescribed size. A control parameter β that plays the role of the temperature is initialized. At certain points in the algorithm the temperature is lowered and for this a "cooling rule" has to be prescribed, e.g. one could take $\beta' = \alpha\,\beta$, say with $\alpha = 0,9$. (Of course, more sophisticated cooling rules have been proposed.) At a *fixed* temperature, several trials to decrease the chosen cost function are made. The number L of trials is prescribed (usually in such a way that it is reasonably likely that all neighbours of a configuration are tried). In our case the cost function was chosen as the number of points in IF_3^6 with distance

> 1 to the code. As a stopping rule one takes (a) the cost function has become 0 (a covering has been found) or (b) for h successive temperatures no trial was accepted (the system is "frozen"). It remains to explain how one trial works. First one picks a random codeword and then a random word at distance 1, not already in the code. The code C' is obtained by replacing the codeword by this neighbour. One calculates Δc, the difference in the cost function of C' and C. Then a random generator picks a number y in the interval [0, 1]. (Now the analog with the Metropolis criterion becomes apparent.) The "trial" C' is accepted as new code if and only if $\exp(-\Delta c / \beta) > y$.
Note that if the cost function has decreased, i.e. $\Delta c < 0$, then the transition is always accepted, but if the cost function has increased, then the trial is accepted with probability $\exp(-\Delta c / \beta)$ and this probability becomes smaller as the "temperature" decreases.

As the numbers show, this technique has been remarkably successful for the football pool problem (and several others). The present author was even more pleased by the fact that the combinatorial construction that holds the present record for $\sigma(8, 3)$ was found by staring at the output of one of the simulated annealing algorithms.

REFERENCES

[1] E.H.L. AARTS AND P.J.M VAN LAARHOVEN, *Statistical Cooling: A General Approach to Combinatorial Optimization Problems*, Philips J. of Research, 40 (1985), pp. 193-226.

[2] M.J. ADAMS, *Subcodes and covering radius*, IEEE Trans. Information Theory, IT 32 (1986), pp. 700-701.

[3] A. BLOKHUIS AND C.W.H. LAM, *Coverings by Rook Domains*, J. Combinatorial Theory, A35 (1984), pp. 240-244.

[4] R.A. BRUALDI, V.S. PLESS AND R.M. WILSON, *Short Codes with a Given Covering Radius*, IEEE Trans. Information Theory (to appear).

[5] R.A. BRUALDI AND V.S. PLESS, *On the covering radius of a code and its subcodes*, preprint.

[6] _____, *On the length of codes with a given covering radius*, preprint.

[7] P.B. BUSSCHBACH, M.G.L. GERRETZEN AND H.C.A. VAN TILBORG, *On the Covering Radius of Binary, Linear Codes Meeting the Griesmer Bound*, IEEE Trans. Information Theory, IT 31 (1985), pp. 465-468.

[8] A.R. CALDERBANK, *Covering Radius and the Chromatic Number of Kneser Graphs, J. Combinatorial Theory (to appear).*

[9] A.R. CALDERBANK AND N.J. SLOANE, *Inequalities for covering codes*, preprint.

[10] G.D. COHEN, M.G. KARPOVSKY, H.F. MATTSON, JR. AND J.R. SCHATZ, *Covering Radius-Survey and Recent Results*, IEEE Trans. Information Theory, IT 31 (1985), pp. 328-343.

[11] G.D. COHEN, A.C. LOBSTEIN AND N.J.A. SLOANE, *Further Results on the Covering Radius of Codes*, IEEE Trans. Information Theory, IT 32 (1986), pp. 680-694.

[12] D.E. DOWNEY AND N.J.A. SLOANE, *The covering radius of cyclic codes of length up to 31*, IEEE Trans. Information Theory, IT 31 (1985), pp. 446-447.

[13] H. FERNANDES AND E. RECHTSCHAFFEN, *The Football Pool Problem for 7 and 8 Matches*, J. Combinatorial Theory, A35 (1985), pp. 109-114.

[14] R.L. GRAHAM AND N.J.A. SLOANE, *On the Covering Radius of Codes*, IEEE Trans. Information Theory, IT 31 (1985), pp. 385-401.

[15] M.W. VAN DER HAM, *Simulated Annealing applied in Coding Theory*, Master's Thesis, Eindhoven University of Technology, (1988).

[16] H.O. HÄMÄLÄINEN, *Upper bounds for football pool problems and mixed covering codes*, preprint.

[17] I.S. HONKALA, *Lower Bounds for Binary Covering Codes*, IEEE Trans. Information Theory, IT 34 (1988), pp. 326-329.

[18] I.S. HONKALA AND H.O. HÄMÄLÄINEN, *A New Family of Covering Codes*, IEEE Trans. Information Theory (to appear).

[19] H. JANWA, *Some New Upper Bounds on the Covering Radius of Codes*, IEEE Trans. Information Theory (to appear).

[20] H. JANWA AND H.F. MATTSON, JR., *Covering Radii of Even Subcodes of t-dense Codes*, Proc. AAECC 3, Grenoble, Lect. Notes Comput. Sci. 229 (1986), pp. 120-130.

[21] H.J.L. KAMPS AND J.H. VAN LINT, *The Football Pool Problem for 5 Matches*, J. Combinatorial Theory, A3 (1967), pp. 315-325.

[22] _____, *A covering problem*, Colloquia Mathematica Societatis János Bolyai, 4 (1970), pp. 679-685.

[23] K.E. KILBY AND N.J.A. SLOANE, *On the Covering Radius Problem for Codes I, II*, SIAM J. Algebraic and Discrete Methods, 8 (1987), pp. 604-627.

[24] S. KIRKPATRICK, C.D. GELATT, JR. AND M.P. VECHI, *Optimization by Simulated Annealing*, Science 220 (1983), pp. 671-680.

[25] P.J.M. VAN LAARHOVEN AND E.H.L. AARTS, *Simulated Annealing: Theory and Applications*, D. Reidel Publishing Company, Kluwer Academic Publishers, Dordrecht, The Netherlands, (1987).

[26] P.J.M. VAN LAARHOVEN, E.H.L. AARTS, J.H. VAN LINT AND L.T. WILLE, *New Upper Bounds for the Football Pool Problem for 6, 7 and 8 Matches*, J. Combinatorial Theory (to appear).

[27] J.H. VAN LINT, JR, *Covering Radius Problems*, Master's Thesis, Eindhoven University of Technology, (1988).

[28] H.F. MATTSON, JR., *An Improved Upper Bound on Covering Radius*, Lecture Notes in Computer Science, 228 (1986), pp. 90-106.

[29] N. METROPOLIS, A. ROSENBLUTH, M. ROSENBLUTH, A. TELLER AND E. TELLER, *Equation of State Calculations by Fast Computer Machines*, J. of Chem. Physics, 21 (19853), pp. 1087-1092.

[30] J. PACH AND J. SPENCER, *Explicit codes with low covering radius*, IEEE Trans. Information Theory (to appear).

[31] J. SIMONIS, *The minimal covering radius t*[15, 6] *of a 6-dimensional binary linear code of length 15 is equal to 4*, IEEE Trans. Information Theory (to appear).

[32] A.N. SKOROBOGATOV, *On the covering radius of BCH codes*, Proc. Third Int. Workshop on Information Theory, Sochi (1987), pp. 308-309.

[33] N.J.A. SLOANE, *A new approach to the covering radius of codes*, J. Combinatorial Theory, A24 (1986), pp. 61-86.

[34] E. VELIKOVA, *Bounds on covering radius of linear codes*, Comptes Rendus de l'Academie bulgare des Sciences, **41** (1988), pp. 13-16.

[35] E.W. WEBER, *On the Football Pool Problem for 6 Matches: A New Upper Bound*, J. Combinatorial Theory, A35 (1983), pp. 109-114.

[36] G.J.M. VAN WEE, *Improved Sphere Bounds on the Covering Radius of Codes*, IEEE Trans. Information Theory, IT 34 (1988), pp. 237-245.

[37] L.T. WILLE, *The Football Pool Problem for 6 Matches: A New Upper Bound Obtained by Simulated Annealing*, J. Combinatorial Theory, A45 (1987), pp. 171-177.