

\*\*\*\*\*The following questions are concerned with Chapter 1 of the notes - Basic Coding Theory.\*\*\*\*\*

- 1 Let  $C = \{00101, 11011, 10100, 10010\} \subseteq \{0, 1\}^5$ . Find  $d(C)$ . Give examples of words that do or do not have a unique nearest neighbour in  $C$ .
- S1 We find that Hamming distances between distinct codewords are 2 or 4. So  $d(C) = 2$ . If the codewords are  $c_1, c_2, c_3, c_4$  in order, and  $w_1 = 10101$ , then  $d(c_1, w_1) = 1 = d(c_3, w_1)$ , so  $w_1$  does not have a unique nearest neighbour. But if  $w_2 = 11111$ , then  $d(c_2, w_2) = 1$  but  $d(c_i, w_2) = 3$  for all the other codewords, so  $c_2$  is  $w_2$ 's unique nearest neighbour.  $\triangle$
- 2 Consider the words GOTHs, HAREs, HATEs, MAREs, MARKs, MATEs, MATHs, MATEY, MITEs, MOTHS, MYTHs, and RITEs. Let these be the codewords of the  $(n, M_1, d_1)$  code  $C_1 \subseteq \{A, B, \dots, Z\}^5$ .
  - a) For each of the words PARKs, GOALs and DATEs, find its nearest neighbour(s) in  $C_1$ .
  - b) Find  $n$ ,  $M_1$  and  $d_1$ . Now find a  $(n, M_2, d_2)$  code  $C_2 \subseteq C_1$  such that  $d(C_2) = 2$  and  $|M_2| \geq 6$ .
  - c) Find three codewords  $x, y$ , and  $z$  in  $C_1$  such that  $d(x, y) = d(x, z) + d(z, y)$ .
  - d) Find three codewords  $x, y$ , and  $z$  in  $C_1$  such that  $d(x, y) < d(x, z) + d(z, y)$ .
- S2 Don't work out 72 Hamming distances! For b), try drawing a graph with a vertex for each codeword, joined by an edge if and only if  $d(c_1, c_2) = 1$ .
  - a) PARKs has n-n MARKs; their Hamming distance is 1. GOALs has n-n GOTHs; their Hamming distance is 2. DATEs has 2 n-ns, HATEs and MATEs, at Hamming distance 1.
  - b)  $n = 5$ ,  $M_1 = 12$ ,  $D_1 = 1$ . There are many options for  $C_2$ , for example  $C_2 = \{ \text{MATEs, RITEs, HAREs, MARKs, MYTHs, GOTHs} \}$  or  $C_2 = \{ \text{MATHs, GOTHs, MATEY, MITEs, HATEs, MARKs} \}$
  - c) For example:  $x = \text{RITEs}$ ,  $y = \text{MATEs}$ ,  $z = \text{MITEs}$ , giving  $2 = 1 + 1$ .
  - d)  $x = \text{MOTHS}$ ,  $y = \text{MYTHs}$ ,  $z = \text{MATHs}$ , giving  $1 < 1 + 1$   $\triangle$
- 3 Let  $C = \{01010, 10101, 11000, 11111\} \subseteq \{0, 1\}^5$ . Find  $d(C)$ . How many errors can  $C$  detect? and how many can it correct?
- S3 By checking  $\binom{4}{2} = 6$  distances,  $d(C) = 2$ . So the code can detect a single symbol error, but not reliably correct any.  $\triangle$
- 4 Let  $C = \{01234, 12340, 23401, 34012, 40123\} \subseteq \{0, 1, 2, 3, 4\}^5$ . Find  $d(C)$ . How many errors can  $C$  detect? and how many can it correct?
- S4 Clearly  $d(C) = 5$ , so  $C$  detects up to 4 symbol-errors, and corrects up to 2.  $\triangle$
- 5 For fixed  $n \geq 1$ , how many binary  $(n, 2, n)$  codes are there?
- S5 Since the minimum difference  $d(C) = n$ , the block length, the two words in the code must differ at every position; for example, with  $n = 5$ , one code is  $C = \{11010, 00101\}$ . There are  $2^n$  options for the first codeword, and then the second is fixed. But a code is a set, so the order doesn't matter:  $C = \{00101, 11010\}$  is the same code. So divide by 2; there are  $2^{n-1}$ .  $\triangle$
- 6 Let  $C$  be an  $(n, M, d)$  code with  $n \geq d \geq 2$ 
  - a) Fix  $j$  with  $1 \leq j \leq n$  and form  $C_1$  by deleting the  $j$ th entry from each word in  $C$ . Show that  $C_1$  is a  $(n-1, M, d)$  or  $(n-1, M, d-1)$  code.
  - b) Form  $C_2$  by deleting the last  $m$  entries of each word in  $C$ . What can we say about the parameters of  $C_2$  if  $m < d$ ? How about if  $m \geq d$ ?
- S6 a) Since we are forgetting the  $j$ -th entry of each word of  $C$  the resulting words will have  $n-1$  symbols, so the block length of  $C_1$  is  $n-1$ . In order to see that  $C_1$  also has  $M$  code words we

have to check that any two distinct words  $x$  and  $y$  of  $C$  give words in  $C_1$  that are distinct as well. But  $d(x, y) \geq d(C) = d \geq 2$  so that  $x$  and  $y$  differ in at least one position other than the  $j$ -th and the resulting words in  $C_1$  will still be distinct. Finally, if  $x_1$  denotes the word in  $C_1$  obtained from  $x$  in  $C$ , then  $d(x_1, y_1) = d(x, y)$  if  $x$  and  $y$  have the same symbol in the  $j$ -th position, and  $d(x_1, y_1) = d(x, y) - 1$  if  $x$  and  $y$  do not have the same symbol in the  $j$ -th position. Now  $d = d(C)$  is the minimum value that  $d(x, y)$  can attain, and  $d(C_1)$  is the minimum value that  $d(x_1, y_1)$  can attain, when  $x \neq y$  are in  $C$ . So  $d(C_1)$  could either be the same as  $d(C)$ , or go down by 1. Therefore  $d(C_1) = d$  or  $d - 1$ .

[Both possibilities can occur. If  $C = \{000, 110\}$ , which is a  $(3, 2, 2)$  code, then  $d(C_1) = 2$  if  $j = 3$  but  $d(C_1) = 1$  if  $j = 1$  or  $2$ .]

b)  $C_2$  has block length  $n - m$ . Any two codewords of  $C$  differ in at least  $d$  places, so removing  $m < d$  places does not remove all differences; the two shortened codewords remain distinct. So  $|C_2| = |C| = M$ . (If, instead,  $m$  were  $\geq d$ , then two words might become identical, so  $|C'|$  could be  $\leq M$ .) We remove up to  $m$  differences from any pair of words. So, if  $m < d$ , then  $d - m \leq d(C_2) \leq d$ . However, if  $m \geq d$ , and two close codewords became identical, the minimum distance could actually go up. For example, let  $C = \{00000, 11111, 11122\} \subseteq \{0, 1, 2\}^5$ , so  $d(C) = 2$ . Taking  $m = 2$  and deleting the last two positions gives  $C_2 = \{000, 111\}$ , with  $d(C_2) = 3$ .  $\triangle$

- 7** A binary code with block length 4 is transmitted over a channel such that  $P(1 \text{ received} \mid 0 \text{ sent}) = 0.1$  and  $P(0 \text{ received} \mid 1 \text{ sent}) = 0.05$ . Is this channel symmetric? If 0001 is sent what is the chance that 0110 is received?
- S7** This channel is not symmetric because the chance of changing a 0 into 1 is different from the chance of changing a 1 into 0. The chance that a 0 sent is received as 0 is 0.9 and that the chance that a 1 sent is received as 1 is 0.95. If  $(0, 0, 0, 1)$  is transmitted, then the first 0 must remain unchanged, the next two 0's must be changed, and the 1 must also be changed. So the chance of this happening is  $(0.9) \times (0.1)^2 \times (0.05) = 0.00045$ .  $\triangle$
- 8** Consider the code  $C = \{c_1, c_2, c_3\} = \{000000, 110000, 111111\} \subseteq \{0, 1\}^6$ , and the words  $w_1 = 010100$ ,  $w_2 = 111100$ ,  $w_3 = 110100$ ,  $w_4 = 111110$ .
- a) Perform nearest neighbour decoding for each  $w_i$ . When is there no unique nearest neighbour?
- b)  $C$  is sent over a binary symmetric channel with symbol-error probability  $p$ . For each  $1 \leq i \leq 3$  and  $1 \leq j \leq 4$ , find  $P(w_j \text{ received} \mid c_i \text{ sent})$ .
- S8** a) We write  $d(x, c)$ , where  $c$  is in  $C$  and  $x$  runs through the four given words, in a table.

	$(0, 0, 0, 0, 0, 0)$	$(1, 1, 0, 0, 0, 0)$	$(1, 1, 1, 1, 1, 1)$
$(0, 1, 0, 1, 0, 0)$	2	2	4
$(1, 1, 1, 1, 0, 0)$	4	2	2
$(1, 1, 0, 1, 0, 0)$	3	1	3
$(1, 1, 1, 1, 1, 0)$	5	3	1

From the rows it is then clear that there is no unique nearest neighbour in the first two cases, and there is a unique nearest neighbour in the last two cases. So  $(1, 1, 0, 1, 0, 0)$  decodes to its unique nearest neighbour  $(1, 1, 0, 0, 0, 0)$  and  $(1, 1, 1, 1, 1, 0)$  decodes to its unique nearest neighbour  $(1, 1, 1, 1, 1, 1)$ . For the first two words we have a choice. We can decode  $(0, 1, 0, 1, 0, 0)$  as either  $(0, 0, 0, 0, 0, 0)$  or as  $(1, 1, 0, 0, 0, 0)$ , and we can decode  $(1, 1, 1, 1, 0, 0)$  as either  $(1, 1, 0, 0, 0, 0)$  or as  $(1, 1, 1, 1, 1, 1)$ .

b) The following table gives the chance that  $x_i$  (in first column) is received when  $c_j$  (in top row) was sent:

	000000	110000	111111
010100	$p^2(1-p)^4$	$p^2(1-p)^4$	$p^4(1-p)^2$
111100	$p^4(1-p)^2$	$p^2(1-p)^4$	$p^2(1-p)^4$
110100	$p^3(1-p)^3$	$p(1-p)^5$	$p^3(1-p)^3$
111110	$p^5(1-p)$	$p^3(1-p)^3$	$p(1-p)^5$

△

- 9** For the binary code  $C = \{0000, 1000, 1111\}$ , the codeword 1111 is transmitted over a binary symmetric channel with symbol-error probability  $p = 0.1$ . We decode a received word to its unique nearest neighbour if it has one; otherwise we do not decode. What is the chance that the received word is decoded correctly? Incorrectly?
- S9** This is a long but straight-forward task. Make a table showing the distance from every possible received word  $y$  (in first column) to each codeword  $c$  (in top row). It is then easy to see which received word has which unique nearest neighbour, or none. Next find the chance of receiving each word  $y$  if in fact 1111 had been sent. And finally we added up the probabilities. If  $p = 0.1$ , chance of decoding correctly (to 1111) is  $(1-p)^4 + 4(1-p)^3p = 0.9477$ . Chance of decoding incorrectly is  $3(1-p)^2p^2 + 4(1-p)p^3 + p^4 = 0.028$  Chance of not decoding at all is  $3(1-p)^2p^2 = 0.0243$ . △
- 10** Consider the binary code  $C = \{000, 111\}$ . Suppose the codewords are transmitted over a binary symmetric channel with symbol-error probability  $p$ . Consider the following strategies:
- (i) Complete decoding using nearest neighbour decoding.
  - (ii) Accepting a received word if it is in  $C$  but asking for retransmission otherwise.
- For each strategy find the chance that, when we send 000, it is decoded correctly, perhaps after several transmissions. If  $p = 0.1$ , which method is more reliable? Should we therefore use this method?
- S10** Note that  $C$  is perfect, and that  $(0, 0, 0)$ ,  $(1, 0, 0)$ ,  $(0, 1, 0)$  and  $(0, 0, 1)$  have  $(0, 0, 0)$  as unique nearest neighbour, and that  $(1, 1, 1)$ ,  $(0, 1, 1)$ ,  $(1, 0, 1)$  and  $(1, 1, 0)$  have  $(1, 1, 1)$  as unique nearest neighbour. For method (i) we decode the received word correctly if it has  $(0, 0, 0)$  as nearest neighbour, i.e., has either zero or one 1's in it. The chance for this is  $(1-p)^3$  (for receiving  $(0, 0, 0)$ ) and  $p(1-p)^2$  for each of the words with one 1 in it. In total this gives  $(1-p)^3 + 3p(1-p)^2$  as the chance of decoding correctly. Since we always decode (i.e., the decoding is complete) it goes wrong in all other cases, so the chance of decoding incorrectly is  $1 - (1-p)^3 - 3p(1-p)^2$ . For (ii) it helps to draw a tree diagram. We decode correctly if we receive  $(0, 0, 0)$  directly, or we receive a non-code word and ask for retransmission  $n$  times ( $n \geq 1$ ) in a row, and then receive  $(0, 0, 0)$  the  $(n+1)$ st time. The chance  $s$  of having to ask for retransmission when we receive a word is the chance of receiving anything but  $(0, 0, 0)$  and  $(1, 1, 1)$ , i.e.,  $1 - (1-p)^3 - p^3$ . [Note that  $s$  is also equal to  $3p(1-p)^2 + 3p^2(1-p)$ , the chance of receiving one of the six non-code words.] So the chance of decoding correctly is  $(1-p)^3 + s(1-p)^3 + s^2(1-p)^3 + \dots = (1-p)^3/(1-s)$  since  $0 < s < 1$ . Similarly we decode incorrectly only if the first time that we receive a codeword that codeword is  $(1, 1, 1)$ . This has chance  $p^3 + sp^3 + s^2p^3 + \dots = p^3/(1-s)$ , a geometric sum. Putting  $p = 0.1$  gives the chance of correct decoding at 0.972 for (i), 0.9986 for (ii). So (ii) is more reliable - but also slower and more costly. △

- 11** The ternary code  $C = \{01, 02, 20\}$  is transmitted over a ternary symmetric channel with error probability  $p = 0.02$ . We decode received words as the nearest neighbour if that is unique, and ask for retransmission otherwise.
- If 02 is sent, what is the chance that it is received as a word in the code?
  - If 01 is sent, what is the chance that we ask for retransmission?
- (Hint for part b): first find which received words do not have a unique nearest neighbour.)
- S11** a) If  $(0, 2)$  is sent then the chance that it is received as  $(0, 2)$  is  $(1 - p)^2$ , the chance that it is received as  $(0, 1)$  is  $(1 - p)p/2$  (because a *specific* error must occur in the second position, and the chance of that is  $p/2$  since the code is ternary), and the chance that it is received as  $(2, 0)$  is  $(p/2)^2$  (because now a specific error must occur in both positions). So the chance that the received word is any of the three code words is the sum  $(1 - p)^2 + (1 - p)p/2 + (p/2)^2 = 0.9703$ .
- b) We only ask for retransmission when we receive a word that does not have a unique nearest neighbour in  $C$ , so we have to find those words first. There are nine possible received words (the elements of  $\{0, 1, 2\}^2$ ), and checking which ones do have a unique nearest neighbour and which do not shows that  $(0, 1)$  is the unique nearest neighbour of  $(0, 1)$  and  $(1, 1)$ ;  $(0, 2)$  is the unique nearest neighbour of  $(0, 2)$  and  $(1, 2)$ ;  $(2, 0)$  is the unique nearest neighbour of  $(1, 0)$  and  $(2, 0)$ ; and that the remaining words  $(0, 0)$ ,  $(2, 1)$  and  $(2, 2)$  do not have a unique nearest neighbour in  $C$ . If  $(0, 1)$  is sent then the chance that  $(0, 0)$  is received is  $(1 - p)p/2$ ; the chance that  $(2, 1)$  is received is also  $(1 - p)p/2$ ; and the chance that  $(2, 2)$  is received is  $(p/2)^2$ . So the chance that any of those three words is received (so that we ask for retransmission) equals  $(1 - p)p + (p/2)^2 = 0.0197$ .  $\triangle$
- 12** Consider the codes  $C_1 = \{0, 1, 2\}$  and  $C_2 = \{000, 111, 222\} \subseteq \{0, 1, 2\}^3$ , which are sent over a ternary symmetric channel with symbol-error probability  $p$ .
- Find the minimum distances for  $C_1$  and  $C_2$ . How many errors can  $C_1$  and  $C_2$  detect or correct?
  - For  $C_1$  the codeword 0 is sent. What is the chance that the received word is decoded correctly under nearest neighbour decoding?
  - For  $C_2$  the codeword 000 is sent. Determine the chance that the received word is decoded correctly if we do incomplete nearest neighbour decoding, where we only decode a received word  $x$  if  $d(x, c) \leq 1$  for some  $c \in C_2$  and do not do anything otherwise. What is the chance that we do not decode the received word at all?
  - Again, for  $C_2$  the codeword 000 is sent, but now we accept only codewords as received words, and ask for retransmission otherwise. What is the chance that we receive a codeword the first time? What is the chance that we eventually decode the received word correctly, perhaps after several transmissions?
  - Now take  $p = 0.1$  and compare the chance of failure for parts b), c) and d), where failure means that we decode either incorrectly or not at all, even after several transmissions.
- S12** a)  $d(C_1) = 1$ , so  $C_1$  detects 0 errors, corrects 0 errors.  $d(C_2) = 3$ , so  $C_2$  detects 2 errors, corrects 1 error.
- $C_1$  is only decoded correctly if received correctly, so  $1 - p$ .
  - We decode correctly if we receive 000 (chance  $(1 - p)^3$ ), but also if we receive 001 (chance  $\frac{p}{2}(1 - p)^2$ ) and there are six such words, with one symbol error. So the chance of correct decoding is  $(1 - p)^3 + 3(1 - p)^2$ . We do not decode at all if and only if the received word contains a 0, a 1, and a 2. There are six such words and each has chance  $(\frac{p}{2})^2(1 - p)$ , so in all the chance of not decoding is  $\frac{3}{2}p^2(1 - p)$ .
  - The chance of receiving 111 is  $(\frac{p}{2})^3$ , and the same for 222. So the chance of receiving a codeword is  $(1 - p)^3 + \frac{p^3}{4}$ . So the chance of asking for retransmission, because we did not receive a codeword,

is  $r = 1 - \left((1-p)^3 + \frac{p^3}{4}\right)$ . By drawing a tree diagram, we can see that the chance of eventual success is

$$(1-p)^3[1 + r + r^2 + \dots] = \frac{(1-p)^3}{1-r} = \frac{(1-p)^3}{(1-p)^3 + \frac{p^3}{4}}.$$

We know the geometric sum converges because  $r < 1$ .

e) For  $p = 0.1$  the chance of *failure* is: b) 0.1 c) 0.028 d) 0.000343. So the retransmission method is the most reliable, but also the slowest and most expensive.  $\triangle$

**13** Let the code  $C = \{00000, 11111, 22222, 33333\} \subseteq \{0, 1, 2, 3\}^5$  be transmitted over a symmetric 4-ary channel with symbol-error probability  $p = 0.1$ . We assume that each codeword is equally likely to be sent.

a) Find the nearest neighbours of  $w_0 = 00123$  and  $w_1 = 00111$ .

b) If  $c_0 = 00000$  and  $c_1 = 11111$ , find  $\mathbb{P}(w_j \text{ received} \mid c_i \text{ sent})$  for  $0 \leq i, j \leq 1$ .

c) Find  $\mathbb{P}(w_j \text{ received})$  for  $j = 0, 1$ .

d) Find  $\mathbb{P}(c_i \text{ sent} \mid w_j \text{ received})$  for  $0 \leq i, j \leq 1$ .

e) Comment on the following statement: "If 00000 is sent, we are as likely to receive 00111 as 00123. So if we decode 00123 to 00000, we should also decode 00111 to 00000."

f) Do  $\mathbb{P}(00000 \text{ sent} \mid 00111 \text{ received})$  and  $\mathbb{P}(11111 \text{ sent} \mid 00111 \text{ received})$  add up to 1? Should they?

**S13** a) Since we have  $d(w_0, 00000) = 3$ , and  $d(w_0, c) > 3 \forall c \in C, c \neq 00000$ , the nearest neighbour of 00123 is 00000. Similarly, since we have  $d(w_1, 11111) = 2$ , and  $d(w_1, c) > 2 \forall c \in C, c \neq 11111$ , the nearest neighbour of 00111 is 11111.

b) To 3 s.f.,

$$\mathbb{P}(00123 \text{ rec} \mid 00000 \text{ sent}) = \left(\frac{p}{3}\right)^3 (1-p)^2 = 3 \times 10^{-5}$$

$$\mathbb{P}(00123 \text{ rec} \mid 11111 \text{ sent}) = \left(\frac{p}{3}\right)^4 (1-p)^1 = 1.11 \times 10^{-6}$$

$$\mathbb{P}(00111 \text{ rec} \mid 00000 \text{ sent}) = \left(\frac{p}{3}\right)^3 (1-p)^2 = 3 \times 10^{-5}$$

$$\mathbb{P}(00111 \text{ rec} \mid 11111 \text{ sent}) = \left(\frac{p}{3}\right)^2 (1-p)^3 = 8.1 \times 10^{-4}$$

c) Note first that

$$\mathbb{P}(00123 \text{ rec} \mid 11111 \text{ sent}) = \mathbb{P}(00123 \text{ rec} \mid 22222 \text{ sent}) = \mathbb{P}(00123 \text{ rec} \mid 33333 \text{ sent}),$$

and that

$$\mathbb{P}(00111 \text{ rec} \mid 22222 \text{ sent}) = \mathbb{P}(00111 \text{ rec} \mid 33333 \text{ sent}) = \left(\frac{p}{3}\right)^5 = 4.12 \times 10^{-8}.$$

Then,

$$\begin{aligned}
 \mathbb{P}(00123 \text{ rec}) &= \sum_{i=0}^3 \mathbb{P}(c_i \text{ sent}) \mathbb{P}(00123 \text{ rec} \mid c_i \text{ sent}) \\
 &= \frac{1}{4} \sum_{i=0}^3 \mathbb{P}(00123 \text{ rec} \mid c_i \text{ sent}) \\
 &= \frac{1}{4} \left[ \left(\frac{p}{3}\right)^3 (1-p)^2 + 3 \left(\frac{p}{3}\right)^4 (1-p)^1 \right] \\
 &= 8.33 \times 10^{-6}
 \end{aligned}$$

and similarly,

$$\begin{aligned}
 \mathbb{P}(00111 \text{ rec}) &= \frac{1}{4} \left[ \left(\frac{p}{3}\right)^3 (1-p)^2 + \left(\frac{p}{3}\right)^2 (1-p)^3 + 2 \left(\frac{p}{3}\right)^5 \right] \\
 &= 2.10 \times 10^{-4}
 \end{aligned}$$

d)

$$\begin{aligned}
 \mathbb{P}(00000 \text{ sent} \mid 00123 \text{ rec}) &= \frac{\mathbb{P}(00000 \text{ sent}) \mathbb{P}(00123 \text{ rec} \mid 00000 \text{ sent})}{\mathbb{P}(00123 \text{ rec})} \\
 &= \frac{\frac{1}{4} \times 3 \times 10^{-5}}{8.33 \times 10^{-6}} = 0.900
 \end{aligned}$$

and similarly

$$\begin{aligned}
 \mathbb{P}(11111 \text{ sent} \mid 00123 \text{ rec}) &= \frac{\frac{1}{4} \times 1.11 \times 10^{-6}}{8.33 \times 10^{-6}} = 0.0333, \\
 \mathbb{P}(00000 \text{ sent} \mid 00111 \text{ rec}) &= \frac{\frac{1}{4} \times 3 \times 10^{-5}}{2.10 \times 10^{-4}} = 0.0357, \\
 \mathbb{P}(11111 \text{ sent} \mid 00111 \text{ rec}) &= \frac{\frac{1}{4} \times 8.1 \times 10^{-4}}{2.10 \times 10^{-4}} = 0.964,
 \end{aligned}$$

e) "If 00000 is sent, we are as likely to receive 00111 as 00123." is correct.

But "So if we decode 00123 to 00000, we should also decode 00111 to 00000." is wrong.

To decode 00111, we compare  $\mathbb{P}(00000 \text{ sent} \mid 00111 \text{ rec})$  to  $\mathbb{P}(11111 \text{ sent} \mid 00111 \text{ rec})$ . Since the second is larger, 11111 is more likely to have been sent. The underlying reason is that 00111 is more likely overall to be received than 00123. So the equal likelihood of receiving 00111 or 00123 if 00000 was sent, makes a smaller fraction of the total.

f) If you calculate to at least 5 decimal places,  $\mathbb{P}(00000 \text{ sent} \mid 00111 \text{ rec})$  and  $\mathbb{P}(11111 \text{ sent} \mid 00111 \text{ rec})$  add up to just less than 1. This is because of the very small possibilities that, if 00111 was received, either 22222 or 33333 was sent.  $\triangle$

**14** Consider words of length 3 made using the alphabet  $A = \{0, 1, \dots, q-1\}$  where  $q \geq 3$ . Describe  $S(000, r)$  for  $r = 0, 1$  and 2. How many elements are there in each? Do those sets look like spheres if we identify the elements in  $A$  with  $0, 1, \dots, q-1$  in  $\mathbb{R}$ , and view all words in  $\mathbb{R}^3$ ?

**S14**  $S((0, 0, 0), 0) = \{(0, 0, 0)\}$ .  $S((0, 0, 0), 1)$  is all words containing at least 2 zeros. There are  $1+3(q-1)$  of these.  $S((0, 0, 0), 2)$  is all words containing at least 1 zero. There are  $1+3(q-1)+3(q-1)^2$

of these. (You can use Lemma 1.7.)  $S((0, 0, 0), 3) = A^3$ . If we see these words as vectors within  $\mathbb{R}^3$ , then  $S((0, 0, 0), 0)$  is the origin,  $S((0, 0, 0), 1)$  is integer points on all three axes, from 0 to  $q - 1$ , and  $S((0, 0, 0), 2)$  is integer points within three squares of size  $(q - 1) \times (q - 1)$ , contained within the  $xy$ ,  $yz$  and  $zx$  planes respectively, and intersecting in pairs along their edges (the axes). (A picture, or model, would help!) These do not look anything like 'ordinary' spheres, because the Hamming distance is so different from Euclidean distance.  $\triangle$

**15** Let  $C$  be a ternary  $(4, 9, 3)$ -code. Show that  $C$  is perfect.

**S15**  $C$  lives in  $\{0, 1, 2\}^4$ , which contains  $3^4 = 81$  words, and  $|C| = 9$ . Since  $d = 3$  we know that spheres of radius 1 round codewords are disjoint, and by Lemma 1.7,  $|S(c, 1)| = 1 + 4(3 - 1)^1 = 9$ . But then these 9 spheres together cover  $9 \times 9 = 81$  words, so the whole space, as required.  $\triangle$

**16** Let  $C$  be an  $(n, M, 2t)$  code with  $M > 1$ . (In other words,  $d(C)$  is even).

- Given code words  $x$  and  $y$  such that  $d(x, y) = 2t$ , find a word  $z$  not in the code such that  $d(x, z) = d(y, z) = t$ .
- Can  $z$  be in some  $S(u, r)$  with  $u$  in the code and  $r < t$ ?
- Conclude that  $C$  cannot be a perfect code.

**S16** a) You can make  $z$  as follows. Find the  $2t$  places where  $x$  and  $y$  differ. Now start with  $x$ , and in the first  $t$  of these places, change it to match  $y$ .

b) No. If  $z$  were to lie in some such  $S(u, r)$  then  $d(u, z) \leq r < t$  so that by the triangle inequality  $d(u, x) \leq d(u, z) + d(z, x) < 2t = d(C)$ , a contradiction.

c) So, for  $r < t$   $z$  is not in any  $S(c, r)$ ,  $c \in C$ . But for  $r \geq t$ ,  $z$  is in both  $S(x, r)$  and  $S(y, r)$ . So for no value of  $r$  can the spheres  $S(c, r)$  partition  $A^n$ .  $\triangle$