Enhancement of Student Learning in the Lecture Theatre by means of a Radio Frequency Feedback System

by
Martin Kenward Gilbert

A thesis submitted in partial
fulfilment for the degree of
Doctor of Philosophy

University of London
1999

Department of Physics and Astronomy
University College London

UCL

ProQuest Number: 10797850

ProQuest 10797850

Abstract

The principal formal teaching mechanism of universities is the lecture - a cost efficient format where hundreds of students may be taught by a single lecturer. Lectures' learning outcome is less certain; the lecturer has little ability to appraise the nature of the audience's understanding. The introduction of an electronic communication path from students to the lecturer mitigates this disjuncture. This communication path has been introduced and was found to be extensively used during a series of lectures, with each student provided with a handset that is reliable, inexpensive and portable. Upon these handsets are buttons; the data is transmitted at radio frequencies to the lecturer, aggregated and displayed graphically for quick assimilation.

It has been recognised that this communication path permits the direct *measurement* of student's educational behaviour without disturbing the lecture itself. Preliminary results indicating the value of this research methodology have been obtained. A significant correlation was identified between the percentage of questions during the lecture students answered correctly and their previous year's overall academic mark. A correlation was identified between the number of times students initiate communication with the lecturer and the number of questions during the lecture they answered correctly. Mixed evidence was found regarding a possible correlation between the evidence of satisfaction with their learning students provided using the system and the number of questions during the lecture they answered correctly.

The introduction of an electronic communication path into lectures has proved to be an innovation deserving of further research and wider introduction into teaching practice.

Contents

List of figures

List of tables

Acknowledgements

Thanks to everyone associated with the Molecular Physics laboratory - in particular Jon Gingell and Nyk Jones - for both help in this research and much enjoyable socialising. Thanks also to Michael Worton, Brigitte Picot and everyone else involved in finding me the space and time in which to work when it was critical, and Nigel Mason for finding me more when things got *really* hairy. My thanks to Brian Fantini for his invaluable technical assistance during this project.

To all my Guildford friends - Jason Hemingway, Simon Handby, Alan Harding, Howard Thain and everyone else - thanks for providing me with a well-spring of good company and relaxation over numerous years.

Thanks to Stuart Sullivan for being a great housemate and friend - a rock about which I could swirl.

Mum and Dad I owe thanks on more levels that can be expressed - but I am profoundly grateful to have had so much love, conversation and fun with two of the most special people it has been my pleasure to know. My thanks also to Lucy and Paul; Isabel and Len; Fay, Badger, Samantha, Eleanor and Zöe - I am lucky to have a family that is so loving and interesting. I am grateful to Fudge and Biscuit for many a greeting "meow" and purring nudge.

My love and unbounded thanks to Katherine Brown - friend, girlfriend, soul mate.

Finally, my thanks to Eric Ash for being, simultaneously: benign and perceptive, forthright and tactful, supportive and critical. I have been superbly supervised - but more than that, I have had the pleasure of the company of a man to whose qualities I can only hope to aspire.

"It seems to me to be a superlative thing to know the explanation of everything, why it comes to be, why it perishes, why it is"   *Socrates*

Chapter 1   Introduction

"The wide world is all about you; you can fence yourself in, but you cannot for ever fence it out"   J. R. R. Tolkien, 'The Lord of the Rings'

An academic, somehow transported in time from the turn of the century to the current day, could walk into a lecture theatre filled with eager students and fully understand the teaching role he was about to play. Some details of the mechanics of the lecture, such as the use of overhead projectors, would require explanation but the basic configuration has remained the same: a single lecturer delivering what is fundamentally a monologue to an audience. Once he (as it was then) opened his mouth, however, the students would immediately recognise that the research findings being presented were not so such much outdated as antique. Those hundred years have revolutionised knowledge - and even introduced new, radical courses such as electrical engineering. Yet the predominant formal teaching format in all but the most affluent of universities, the lecture, has barely changed. This thesis will demonstrate the means to effect a substantive improvement to the lecture.

It is not hard to understand why lecturing remains the primary form of teaching at almost all universities, as it is a cost-effective means to ensure that all students have been introduced to the material upon which they will be examined, requiring as little as hundredths of a lecturer hour per student. One can be less certain of its learning outcome. "When measures of knowledge are used, the lecture tends to be as efficient as other methods. Alternatively, in those experiments involving measures of retention of information after the end of a course, measures of transfer of knowledge to new situations, or measures of problem solving, thinking, or attitude change, or motivation for further learning, the results tend to show differences favouring discussion methods over lecture" [McKE].[1] Lectures can amount to little more than a curriculum setting exercise, with *understanding* occurring elsewhere (such as during private study).

---

[1] references in this thesis are indicated by [ABCD], where 'abcd' are the first 4 letters of the author's name; where there are repeated references to an extended text (such as a book) the relevant section is indicated by [ABCD pp.1-10]

A lecturer teaching a small number of students can engage in dialogue with them and actively explore and expand their understanding - as happens during tutorials. However, this interaction becomes difficult, or even impossible, when teaching larger classes. The lecturer has to resort to delivering the material largely as a monologue, without the ability to check on the progress of understanding of the students; the students are constrained to a largely passive role, which can easily lead to their disengagement from the lecture. Although "definitions of 'a large class' are somewhat arbitrary" [WARD] a useful operating definition is where these individual student-lecturer interactions are rare - which is very often the case with an audience of thirty or more.

The number of students attending UK higher education has substantially risen over the last two decades, figure 1.1, whilst the financial resources provided for each of these student's education has declined progressively over the same time period, figure 1.2.[2] This financial pressure has to led to a broadly similar teaching function being delivered to a sharply increased number of students, leading to an inevitable upward pressure in the class sizes attending lectures [GIBB]. Lectures where the audience is small enough for the lecturer to enter into meaningful dialogue with individual students have become rare, particularly so in foundation (first year) courses - where material with the widest applicability is taught. Enhancing the learning outcome from large class lectures, whilst retaining their cost efficiency, would realise significant benefits. Students would receive an improved education from each lecture, raising the possibility of both reducing the number of lecture-hours that each student needs to attend and decreasing the need for ancillary teaching resources (such as well-stocked libraries), with the resources released deployed usefully elsewhere.

---

[2] pressure from increased uptake of higher education is not confined to UK universities and applies in many nations - for example to New Zealand, Canada, Norway [ECONa], Brazil and Mexico [ECONb]
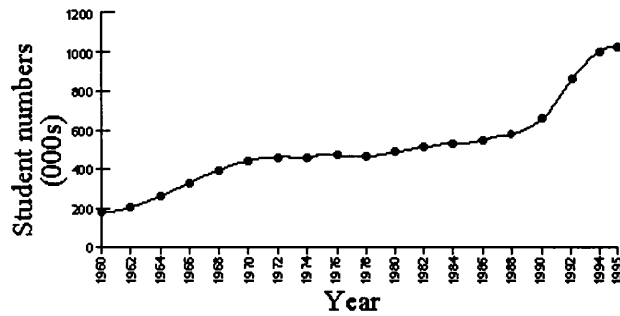
Figure 1.1 -- The number of full-time students attending UK higher education 1960 to 1995, after [DEAR, chart 3.1]
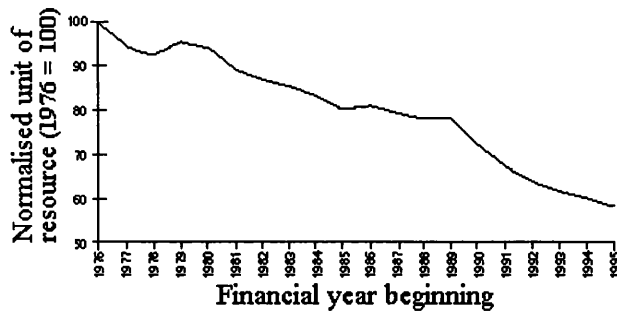
Figure 1.2 -- Public funding per student for the financial years 1976-77 to 1995-6, normalised such that 1976 = 100, after [DEAR, chart 3.16]

An electronic feedback path from the students to the lecturer seems an attractive means to introduce student-lecturer communication into large class lectures, permitting students to convey opinions to the lecturer without the difficulties of other techniques (such as the attendant disturbance and lack of confidentiality of a show of hands). The information provided to the lecturer by this communication path necessarily describes the students' reactions statistically and will allow these opinions to be understood *en masse* by the lecturer. The lecture can then be directed with more accuracy toward those areas which students have yet to understand. It will also provide students with an active role during lecturing, engaging them in the process through their ability to influence its development. This thesis describes the development and initial validification of one such electronic communication path.

To enable the feedback system to be set in a correct context the principal theories of learning are briefly explored. Implementations of technology in an educational

setting are examined, and prior attempts at providing student-lecturer communication paths outlined, to ensure that this research incorporates best practice. These analyses indicate that this electronic communication path should be portable, cheap to introduce, require minimal maintenance and not predicate a particular teaching style. It was determined that the format that best meets these criteria is providing each student with a handset, upon which there are several buttons to be pressed when the student wishes to indicate an opinion. When a button is pressed the data is transmitted via radio to a central receiver, where the data are aggregated and displayed graphically on a computer. These buttons are initially intended to be used in two ways: they can be pressed to respond to multiple choice questions posed by the lecturer; during periods of lecturer discourse they can be pressed to indicate that its speed to "too fast", "all right" or "too slow".

The technical issues surrounding the operation of the feedback system within these guidelines are then examined. The energy required to transmit each button press should be minimised to remove the need for maintenance (such as regular battery replacement). If this energy dissipation were low enough the handsets could then use a photovoltaic power supply; however, in the first instance the system resorts to the use of batteries. The handsets, using CMOS components (the technology of choice for low power electronics), operate on a duty cycle: activating when a button is pressed, transmitting the data and then deactivating. The power supply offered by the use of photovoltaics was provisionally determined to provide an indication of the target energy dissipation per handset operation.

The feedback system is a star shaped network, with many transmitters and a single receiver. Without control over when each handset transmits there is a possibility that two or more handsets' transmissions will occur simultaneously, causing data loss. This probability is evaluated. It is possible to organise the handsets' transmissions, ensuring no overlap between transmissions. However, these methods have the undesirable effect of increasing the energy dissipation per transmission. As the handsets are to be used to gain information regarding the *entire* audiences' opinions (that is, the correct restitution of every transmission is not critical), it was decided

that handset transmissions would not be organised and a probability of data loss from transmission overlap would be tolerated.

A transmission which does not overlap with another is subject to a separate probability of data loss from the effects of random noise. Techniques to reduce this probability are examined, and their impact on the feedback system quantified. The feedback system will use frequency modulation to transmit the data on the radio frequency carrier, and will not deploy error detecting or correcting codes at this stage.

A prototype handset system was constructed and successfully used to conduct initial exploration of the use of the electronic feedback in teaching situations. However, this system had an unacceptably high probability of transmission overlap if more than eight handsets were used at one time, and hence could not be used in large class situations. Additionally, a design flaw caused the handsets' energy dissipation to greatly exceed the design target.

In light of the experience gained in producing the prototype system a large class feedback system was developed. This system was close to the design targets, as it is affordable, can be expanded to 249 handsets and each handset has an energy dissipation per transmission such that its batteries will last for tens of thousands of hours of lecturing.

The results of the exploration of the use of the prototype handset system in teaching situations are presented. It was verified that the system would have utility for an audience over prolonged periods of teaching (the handsets were used by eight students over six consecutive hours of lecturing). In a separate study, a total of 54 second year students were given lectures and a detailed analysis of their use of the system conducted. This analysis shows that, although not intended at the outset, the feedback system opens new vistas for educational research. Finally, future research directions, both technical and educational, for this electronic lecture communication system are identified.

Chapter 2   Education overview

"Be very, very careful. This ride is not for the weak. It's a psycho. A bit like real life. Well, maybe not quite that bad" Jeff Noon, 'Vurt'

The aim of this work is innovation in education. One should not embark on such a venture without seeking guidance from the vast body of educational theory that has evolved. This chapter presents a brief summary of the main ideas that have been advanced, and attempts to gauge how they might guide this project.

Review of the three principal schools of thought concerning individual learning processes - behaviourism, gestalt and constructivism - discovered they illuminate different aspects of the educational process. The system was therefore designed without an attempt to be specifically welded to any single educational direction: it will have equal applicability for lecturers using any, or none, of these ideas.

2.1 Behaviourism

Behaviourism has at its heart the precept that humans can be fully analysed by means of the objective observation of their actions. When a person is presented with a stimulus there will be a response. Desirable responses are encouraged through the process of teaching, by reward and punishment. Abstract terms such as 'consciousness' are not inherent to the explanation of humans' behaviour; this behaviour is a direct product of the external events to which they are exposed.

2.1.1 Classical behaviourism

Classical behaviourism arose from the work of I. Pavlov [PAVL]. In his famous experiment, after a period of a bell being rung when food is presented a dog will salivate simply at the sound of the bell, without feeding. This is an example of 'conditioning'. The unconditioned response to a bell being rung, no response on the part of the dog, is conditioned so the dog associates the noise with food and hence

salivates in response to the noise itself. This work was extended to humans, with classical behaviourists proposing that behaviour is simply the reaction of humans to stimuli in their environment. Learning is the teacher adapting the students' response by rewarding correct ones, and punishing the incorrect - stimulus-response (S-R) links are the essence of education. The actions of humans can be broken into a series of S-R relationships, where a human learns by the formation of a new reaction to events. There are two basic 'laws':

(i) *law of effect* is that when a persons' actions cause them to experience satisfaction in a given situation, where that situation recurs the action is also likely to recur (a S-R link is formed). Further, the greater the pleasure (or, by reflection, the lesser the discomfort) the stronger the link between the situation and action. A later modification to this law was that a learner who is punished for an action will seek alternative S-R links to gain pleasure, whereas reward directly reinforces the link (first proposed in [THOR p.96]).

(ii) *law of exercise* is that, all other things being equal, more frequent repetition of a S-R link results in a more secure connection - that is, practice makes perfect. A later caveat to this law is that the learner must be aware of the successful outcome of his actions (proposed in [THOR, p.95]).

W. B. Watson, an influential researcher in this school, felt humans could be completely explained as biologically determined phenomena. He felt that humans did not have "memory", but skills retained over a period of no practice. He rejected the concept of consciousness as an unprovable assumption. There are no intervening mechanisms such as thought or intention in a S-R link [WATS pp.2-6, 235-248].

2.1.2 Operant behaviourism

B. F. Skinner evolved classical behaviourism into operant behaviourism - explaining learning behaviour in terms of 'operant conditioning' [SKIN]. An operant is a series of voluntary acts on behalf of a learner in response to the environment (for example, referring to a textbook when faced with a problem). These operants can be

reinforced to improve the efficiency of the learner by reward when a correct operant is used in a situation, or punishment when it is incorrect. Skinner also identifies 'respondent' behaviour, which are responses elicited involuntarily (for example, blinking of the eye when an object approaches unexpectedly close). From these concepts Skinner explored an experimental analysis of learning behaviour in presence of reinforcement. He believed if one can successfully analyse the various stimuli producing behaviour one can, by influencing those stimuli appropriately, control behaviour.

2.2 Gestalt

Gestalt psychology is named after the German word 'gestalt' meaning a structure, pattern or configuration. The gestaltists state that the mind can only productively be viewed as a whole [BURT]. The simplicity of the S-R relationship is *over-simplified*, and would be better viewed as S-O-R, where the stimulus occurs, is perceptually organised by the learner and then followed by a response. This can be seen by the fact that a melody is recognisable whichever key the piece is played in - it is the overall organisation of the piece that we respond to, not the precise individual notes. The fundamental activity of perception cannot be divorced from the idea of organisation.

The concept of 'insight' is central to the gestalt school. Insight is a complex reaction to a situation in its entirety suddenly producing a new understanding - a 'flash of inspiration'. The learner suddenly perceives the structural essence of a problem through a full comprehension of the situation posed by it.

There are six laws of gestalt psychology, by which one can understand the organisation of the mind:

(i)　　*Law of figure-ground relations*. Perceptions are organised into coherent 'figures' which stand out from the background as identifiable wholes

(ii)　　*Law of contiguity*. Perceptions are grouped into figures with reference to their proximity in time or space.

(iii)   *Law of similarity.* Similar items are grouped together in patterns.

(iv)   *Law of Prägnanz.* The brain tends to group perceptions into figures of maximum simplicity and balance to give most possible illumination.

(v)   *Law of closure.* Incomplete perceptions are viewed as complete - the learner 'fills in the gaps'.

(vi)   *Law of transposition.* Patterns are recognisable even if their exact nature has changed. A tune can be sung in many keys whilst retaining its identity.

2.3 Constructivism

Constructivism developed in part from the gestalt school, and holds views in radical opposition to behaviourism [LESG][MATT][MART]. Constructivism holds the view that individuals create mental constructs from the information that is given to them. The learner selects important information from the stimuli provided, and rejects that which is felt to be trivial or fallacious. The learner then applies this construct to novel stimuli. If it fails to perform adequately - providing incorrect answers, for example - it is rejected and a new construct is developed in light of the new information. Learning is the process of adjusting conceptions of the world to produce the best possible result. This adjustment occurs within the learner, and cannot necessarily be observed. The learner is an active participant in the learning process, and not a passive recipient.

2.3.1 Classical constructivism

Classical constructivism believes that whilst the learner might have a wide variety of experiences and resultant mental operations, all learners approach a common set of constructs. Leading exponents of this view include J. Piaget, who stated that humans pass through several stages of development from childhood - there is a progression of the mind towards a shared construct [GINS]. This progression, it should be emphasised, is *not necessarily* a linear one - as J. Bruner termed it, one can have a "spiral curriculum" [BRUN]. This takes early learning, moving it upwards in

reformulating constructs through tuition, and then circles back to the previous understanding to permit further progress from further tuition.

2.3.2 Radical constructivism

Radical constructivism takes issue with the idea that learners progress to broadly shared constructs. It views the learner as being unfettered as to what models to draw from his experiences; learners create many different constructs even with the same experiences. Learners should be encouraged to explore their constructs through the active reinterpretation of ideas. According to this school, teaching should centre on students' reappraisal of their constructs with teachers acting as facilitators of many different cognitive changes.

2.4 Implementation

The execution of these ideas in lectures with student-lecturer communication would lead to differing teaching styles. A behaviourist lecturer would be seeking to present the information, and then test the students upon it. Those answering correctly would get a reward - perhaps tokens towards a free coffee (and those responding incorrectly could receive a punishment, such as the requirement to return the tokens). If the lecturer felt too many students answered incorrectly the material would be outlined again, with further tests, until it was felt that the students had developed the correct response. A gestaltist lecturer would present material with reference to its relation to other concepts - its location in the overall picture. The feedback would then be used to examine the nature of the figures the students had developed. Incorrect figures, or a lack of insight into the new figure, would lead to further teaching until correct ones were in place. A constructivist lecturer would first seek to explore what the students' pre-existing constructs were. In light of that information, the lecture would attempt to highlight the failures of these constructs, by presenting contradictory evidence. Following this, the nature of the constructs the students had developed during the lecture would then be examined. If it was felt that these new constructs differed too

strongly from that intended, further contrary facts and greater elucidation of the material would follow.

2.5 False trichotomy?

The three schools of thought outlined are not necessarily mutually exclusive. Indeed, to a certain extent they overlap. The underlying philosophy of the post-teaching testing differs, but feedback is used to test students' understanding of the material presented in every case.

A further difficulty arises from the fact that these educational ideas were developed with reference to *school* students (and animals). These learners are significantly less complex than university students under *all* of the paradigms (fewer S-R links, less operant conditioning, fewer figure-ground relationships or less detailed constructs). This adds another layer of opacity to their application in university teaching.

A typical university provides many different types of learning, from the practical manipulation of experimental apparatus in the sciences, to the use of syntax and grammar in the arts (and occasionally sciences), to highly abstract concepts. Within lectures I do not feel that any one theory of learning is applicable, but rather that there is a spectrum of utility, where different types of learning play varying roles.

Finally, it should be noted that many lecturers do not teach with *any* of the three main paradigms uppermost in their mind. Indeed, some are not aware of them at all. If the feedback system was designed for only one theory, this would necessitate an accompanying explanation of the theory and the manner in which the feedback should be used. This adds complexity to the system's introduction, which would have to be balanced by improved results from the application of that theory. There is no unequivocal evidence to indicate the universal superiority of any one of these ideas.

In light of these difficulties, it was considered desirable to design the feedback system *purely* as a means to improve student-lecturer communication. This communication is not entirely educationally a-theoretical. Behaviourist communication views the student as a passive recipient of knowledge. Feedback from the student is viewed as permitting the examination of their knowledge for possible reinforcement. Constructivism and gestaltism, however, view the learner as active participant: communication is to elicit the learners' internal response to aid the teacher in understanding the nature of their mental models. Nevertheless, learner-teacher communication fundamentally underlies all teaching methods. Without adequate communication of ideas and concepts it is not possible for the student to learn, and the *provision* of the means to effect feedback is theoretically neutral. It is the lecturer's choice how utilise the system - negating any possible 'Not Invented Here' syndrome (lecturers feeling constrained to teach to a particular pedagogy that they do not appreciate, and therefore abandoning the system).

Chapter 3   Educational technology

"now it's all reversed. The media image is the reality, and by comparison day to day life seems to lack excitement. So now day to day life is false, and the media image is true. Sometimes I look around my living room, and the most real thing is the television. It's bright and vivid, and the rest of my life looks drab. So I turn the damn thing off. That does it every time. Get my life back." Michael Crichton, 'Airframe'

New technology has been declaimed as being on the point of revolutionising university education since the late eighteenth century - as Samuel Johnson remarked: "Lectures were once useful; but now, when all can read, and books are so numerous, lectures are unnecessary" [BOSW]. These claims have persisted, with Bill Gates writing "PCs [Personal Computers] can do a better job of supporting varied thinking and learning modes than lecturers and textbooks can" and "the daily habits of students and faculty will change to take advantage of the opportunities interactive networks offer. The many small changes will add up [...] to change significantly the formal processes of education." [GATE] To date, the outcomes of these claims have been somewhat less dramatic.

This chapter will argue that the introduction of some technologies has improved the delivery of university teaching, and will continue to do so. These technologies have supplemented existing teaching formats, for example the use of text in tandem with lecturing; or replaced elements within a teaching format as with the use of overhead projections (OHPs) largely replacing chalk and blackboard within lecture theatres. Other technologies have made less impact, such as the predicted widespread use of television as a medium for the delivery of university teaching [IVES]. A technology is most likely to be successfully introduced if it satisfies a number of criteria. The changes it provides should address specific weaknesses in a teaching situation. The development of educational material using the technology and, critically, its delivery in the educational situation should be as low-cost as possible.[1] The technology

---

[1] the development of the technology *itself* is quite separate from this, and often occurs before any attempt to apply to education

should also offer discernible benefits, either by improving the quality of teaching outcome, or by providing savings in the time and/or cost required for the overall teaching function.

The need to improve the quantity, quality and immediacy of student-lecturer communication in large class lecturing has long been recognised (for example [BLIG pp.99-121]). There have been various previous attempts to introduce technology to encourage this communication without widespread success - although where there has been long-term use significant learning improvement is apparent [POULa]. Possible reasons for this lack of take-up are identified to guide the development of radio frequency lecture feedback. It will be reliable, flexible, portable, inexpensive and requiring minimal maintenance. Further, the physical process of student data transmission will be as confidential as possible, and it will be presented to the lecturer without identification of the originator. The feedback may be initiated by students, so that they can indicate when they are having difficulty in understanding the lecture, or by the lecturer posing multiple choice questions.

3.1 Historical outcomes of educational technology

There have been many attempts to introduce promising technologies into university education, but few successes. Diana Laurillard quotes "a folk wisdom in academic circles that educational technologies come and go, and all the expensive machines end up gathering dust in cupboards." [LAUR p.210] This is somewhat unfair, not least as it implicitly defines educational technologies as "expensive machines", and the subsequent discussion also focuses on newer educational technologies such as audio-visual media and the use of computers in education. It is important to recognise older technologies such as OHPs as providing very useful case studies of success; with a case study of an unsuccessful technology being provided by one form of computer mediated learning.

3.1.1 Case study of a successful technology

One comparatively recent and highly widespread use of an educational technology is that of overhead projection during lecturing - Lewis Elton claims the "two most successful aids in education over the past twenty years have not been television and the computer, although they have had the most publicity, but the overhead projector and the Xerox machine" [ELTO]. However, there is a dearth of study of the teaching effectiveness, and ancillary impact, of lecturing using OHPs as compared to the traditional chalk-on-blackboard delivery. I am forced, therefore, to surmise the attractions that lead to its ubiquity.

Some attractions can perhaps best be understood by observing where OHPs have *not* found widespread use - mathematics teaching, where the use of blackboards remained common. In broad terms, mathematics relies on the comprehension of a series of logical procedures that lead toward a conclusion - each step is *as* important as the next. Therefore, there are fewer opportunities to present key conceptual crystallisations of material. It is important to note that OHPs can be used in exactly the same way as blackboards, with the lecturer writing on overheads during the discourse (and devices such as the continuous feed, roller to roller, transparency even obviates the difficulty of changing single transparency pages). The lack of OHP use in this manner could be due to the difficulty of quickly writing large quantities of clear notation in a font that is small to the lecturer.

Overhead projections in lectures are not often used in this direct blackboard replacement format - more common are individual transparency pages upon which the lecturer has prepared written information. Subjects taught in OHP assisted lectures often do have conceptual crystallisations. For example, in physics laws and concepts can be introduced in a written format, and their effects and manifestations subsequently discussed - the key material can be condensed and presented in bullet point format. The OHP can be used to reproduce, identify and emphasise important material; the lecturer is then freed from the 'mundane' task of writing and can talk about the subject more freely. This condensation could be written on a blackboard,

but there are further additional logistical advantages with the use of prepared transparencies: ensuring the presentation of all the material intended with an ordered delivery, little chance of missing important topics, the writing itself is likely to as legible as possible and, above all, the lecturer is released from the chore and distraction of writing during the lecture. It also reasonably simple to use colour to consistently highlight or identify categories of material.[2] OHPs can additionally be used to present diagrams and pictures without the need to create them during the lecture itself (such pictures commonly also represent a condensation of information). Many OHP lectures also include textbook references and handouts of more detailed material - the latter often photocopied - so much so that Laurillard comments that this "combination of lecture and print has almost become the canonical form of the 'lecture'" [LAUR p.110]. In this case, printed text is used to pass a detailed permanent copy of factually deductive and logical information to the students, removing the need to take copious amounts of notes.

Therefore one can discern that the use of OHPs frees the lecturer, and to a degree students, from the task of writing during the lecture. This allows the lecturer a more wide ranging teaching role - not least by facing the class - and ensures important points are highlighted and can be readily understood by students who are able to concentrate on understanding the material presented. Additionally, once OHPs have been written, they *stay* written: there are time savings to the lecturer over a multi-year timescale.

There are also some disadvantages that are avoided in the use of overhead projection. It is a robust technology with little chance of an unexpected breakdown; once installed the projector and screen cost very little to maintain, and the marginal cost of the creation of each overhead projection is extremely low.

---

[2] although multiple colours of chalk can be used, there is a delay in locating and bringing to bear a new colour, discouraging frequent colour changes

3.1.2 Case study of two less successful technologies

There are few examples of technologies that have been introduced into university education that have *entirely* failed to have an impact. However, many have yet to progress beyond use in specific situations, such as the delivery of teaching via television which is largely confined to distance learning (for example, by the Open University). This is despite initial claims that videotaped lectures would "reduce costs per student-hour, while holding outputs constant" whilst "replacing bad instructors by good ones" [LAYA]. A clue to this medium's lack of penetration is, surprisingly, to be found within the same paper where it recognised that "videotape can never justify its fixed presentation costs unless the equipment is used for at least 15 hours [per year][...] Similarly, whatever the number of hours' usage, it can never justify its master costs unless each course-hour is used by at least two institutions". The authors felt this a surmountable obstacle - but it does require a considerable initial commitment by those two institutions, and the figures also rely on a series of estimations of the costs of producing videotape that were acknowledged to be at the cheaper end of the spectrum of possibilities. Evidently, the commitment felt necessary was not forthcoming.[3]

There is a good example of a thoroughly unsuccessful technology intended for use in schools, Viewdata, principally because of a remarkably candid paper from one of the technology's advocates, Oleg Liber [LIBEa]. Viewdata was the provision of computer mediated teaching materials. A large database of pages, each "containing text and simple mosaic graphics in up to 7 colours, similar to teletext pages" could be accessed via a "basic microcomputer" with a modem connection to ordinary telephone lines; material could be "created and stored locally, but accessed globally" [ibid.]. These pages were to be used principally as a factual information retrieval facility by schools, to which students could refer. The development of this information delivery mechanism was hailed as having the potential to be as important as the development of the printing press [LIBEb].

---

[3] the Open University has been more successful not least because many of the delivery costs of the material are borne by the students themselves - they buy and maintain the television sets. Additionally, television has the singular advantage in distance learning of being able to be received without the students having to congregate

However, fifteen years after its introduction in the early 1980s Viewdata was defunct. It could perhaps be described as a technological solution in search of an educational problem. It had no clearly defined role within its education milieu and solved few pre-existing difficulties in the access to information (the vast majority of schoolchildren can visit libraries); neither was a new, supplementary education modality developed. Indeed, Liber notes that the longest lasting element of the Viewdata network was one that had not been initially envisaged as being important: the later development of pages that permitted interpersonal interaction, as "people preferred to use systems that allowed discussion. Communication is preferable to information" [LIBEa].

Viewdata not only had a lack of clarity in its educational objectives, it also had a high marginal cost per use. Even after the purchase of the necessary technical equipment, schools had to pay the costs of the telephone link to access the pages, and "cuts in education funding led to many [local education] authorities to limit school's telephone budgets, and schools turned their focus on to other aspects of IT [information technology]" [ibid.].

3.1.3 Summary

The reasons for a technology's success, or lack of it, can often be determined after the attempt at introduction. It is more important to seek guidelines in light of past experiences against which a technology can be judged *before* expenditure is incurred.

Laurillard suggests that the technology must have a "learning context" including: "student preparation" making them aware of the objectives of the learning and its place within the wider course, "integration with the rest of the course" to ensure that the technology is not treated as a bolt-on, "inculcating an appropriate conception of learning" and providing "supporting materials"; academic "assessment" should cover the materials introduced by the technology, and the "logistics" must not be neglected [LAUR pp.210-222]. This is most likely to occur where a technology is introduced

to an existing teaching situation, where many of these elements will already be in place. This point is made more forcibly, with reference to a number of educational technology innovations, by Steve Draper, who states that successful outcomes are "where a teacher analysed what was particularly weak in an existing situation and thought of how technology could be used to address that bottleneck" and offers the following prescription: "take a specific teaching and learning situation; identify the main limitation in the current delivery method; design a solution. Leave other things alone" [DRAP].

There is a further practical requirement that should be added to these educational themes: the technology must be robust as its failure could easily forestall learning. There are also economic considerations. The technology should not be prohibitively expensive to acquire, including such externalities as the space and facilities required for its use. It should not need frequent replacement (the acceptable frequency is, of course, dependent on the required capital). However, the recurrent costs of delivery are probably more important. The cost and time needed to develop teaching material for use with the technology should be minimised. The time required to switch between items of teaching whilst the technology is in use should also be as short as possible. Further, the students' use of the technology should not incur significant expense (for example, the technology should require minimal maintenance).

A recent report, Information Technology Assisted Teaching and Learning (ITATL) in UK Higher Education, prepared for the National Committee of Inquiry into Higher Education (the "Dearing Report" [DEAR]) found that these practical and economic considerations were most likely to be met by techniques that are *not* state of the art.[4] "There appears to be a relatively strong case for ITATL investments which involve less complex technologies and applications. Typically these are robust, require low maintenance [and] support and the processes of adoption are simplified. We recommend that funding agencies support the adoption of less complex technologies where demonstrable economies of scale or quantifiable improvements in learning can be anticipated" [DAVI].

---

[4] sadly, however, this conclusion was not included in the Dearing report which concentrated on the possible educational technological use of new computer software in higher education

3.2 Improving lecturing

Lecturing is commonly recognised as being a teaching modality that deserves improvement. There is a need for training for lecturers, so that best practice can be adopted - such as preparing lectures in advance, communicating information efficiently, etc. [BLIG pp.59-152, 180-191]. However, a more fundamental difficulty for lecturers (however well trained) is that a lecture "provides little opportunity for judging audience understanding and reaction accurately" [GREG]. The didactic flow is unidirectional - there is no "conversational framework" within which the teacher and students can interact [LAUR pp.102-105]. Additionally, students have little formal role with the lecture providing "limited opportunity for active student participation" [GREG]. This passivity can all too often degenerate into a total disengagement from the lecture, replaced by thoughts on possibly more agreeable matters.

Recent research into the improvement of lectures has concentrated on replacement of all or parts of the lecture to encourage students to engage in the teaching process. This can be ascribed in part to a tendency for lecturing to be dismissed as "a grossly inefficient way of engaging with academic knowledge" which should provide a "baseline" from which to escape [LAUR pp.107-109]. Many papers propose changing the lecture format by introducing student-student interaction, or 'buzz groups' (for example, [JACK]). The idea is to break the lecturer's discourse into smaller sections, interspersed with periods where students actively discuss amongst themselves problems or issues raised by the lecturer, after which their conclusions are made known (by a show of hands, etc.). The intention is to "break the unavoidable monotony of the passive lecture" [MAZU]. However, whilst providing students with a role (periodically), this idea does not address the central difficulty of the lecturing itself being a monologue, and any difficulties in student conceptions not apprehended until the next buzz group.

However, there is the possibility of engaging students *during* the discourse itself - to use feedback from the audience to the lecturer to pass information on the progress of their understanding, which also provides students with a more active role. Lecture theatre feedback is not a new idea, *inter alia* being identified as a promising concept by Bligh, but has not found widespread use [BLIG pp.99-121].

At a most basic level, student feedback in response to questioning is possible within lecture theatres without technology, by a show of hands. However, this has disadvantages: it requires students to publicly associate themselves with an answer, potentially leading to students feeling exposed to ridicule; it is time consuming to ask questions with more than two possible answers,[5] and it can be difficult to ascertain with accuracy the proportion of students with their hands up.

3.2.1 Prior attempts to deploy lecture feedback

There have been prior attempts to improve upon 'show of hands' feedback. One such technique was the "Feedback Classroom": each seat in a classroom is provided with a "response unit" upon which there are several buttons, figure 3.1, connected via wire to a "teacher control unit" [WEBB][HOLL]. The lecturer pre-selects the correct answer on the teacher's control unit, and then poses a multiple choice question. The students press the button corresponding to their answer. If a student's answer is correct, a small partially screened lamp on the response unit is illuminated; the lecturer has a set of lamps to indicate every student's answer whether correct or not.

This solution has the distinct advantage that students' answers are confidential. However, the system is relatively cumbersome, with many wired connections, requires installation before use and cannot readily be expanded to very large classes as the number of lamps on the teachers console would become unwieldy. Additionally, the lecturer pre-selection of the correct answer is time-consuming. Perhaps for these reasons the Feedback Classroom concept fell into disuse, with no published analysis of its effects on students learning.

---

[5] as a show of hands can only indicate binary "yes/no" responses the lecturer would have to ask "Was this your answer?" for each of the possibilities
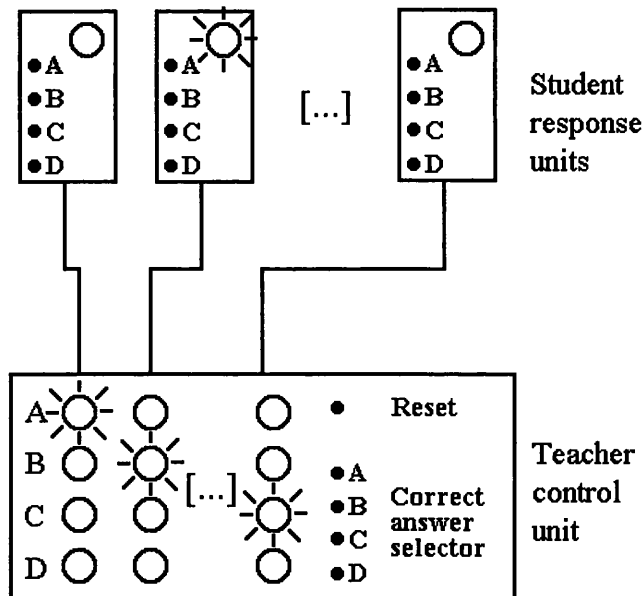
Figure 3.1 -- The "Feedback Classroom"; each student is provided with a response unit for multiple choice questioning. If a student's answer is correct (determined by pre-selection on the teacher's console - 'B' above) a lamp is illuminated. The teacher can see all the answers offered by students, also by lamp illumination. After each question has been answered the teacher resets the system by pressing the appropriate button

Another form of hardwired feedback was introduced successfully and used over many years, "Audience Paced Feedback" (APF), which is essentially similar to the Feedback Classroom, but with simplifications [POULb]. There is no 'correct/incorrect answer' lamp on these units; the system has only a single button on each response unit and questions are phrased so that they can be answered by a simple "yes" or "no". The lecturer can then intersperse periods of feedback between discourse of material. This feedback is used in various ways:

"(a) exploration, to gauge the opinion of the students of their *own* understanding (for example, "Have you understood my arguments regarding this equation?")

(b) verification, to allow the *lecturer* to assess the state of the students comprehension ("Does this apply to high temperatures?")

(c)   interrogation, to test the ability of the students to apply the work to specific situations (for example a multiple choice question [...])

(d)   organization ("Are you ready for me to continue?")" [POULa]

The lecturer can then pose questions, with the students' responses displayed by a dial which indicates the *proportion* of students pressing their button (permitting its use in large class lectures). The lecturer does not need to physically pre-select a correct answer; after each use the system is reset by a button press. The technology is robust and the marginal cost and time needed for its operation by the lecturer is low. Analysis indicates a highly positive educational outcome. The results of approximately 2,500 student-years of APF lecturing was compared to a similar number of student-years of lecturing without feedback (with the traditional lectures in the preceding or succeeding years of APF lectures). The chosen evaluation tool was the examination pass rate of the attending students, as it will provide an indication of the students' learning. APF was found both to significantly increase the pass rate and reduce its year on year variation: students both understand the material better and the level of attainment is more consistent over several years [POULa]. This is an encouraging endorsement that feedback can have a role to positive role to play in lecturing.

However, this feedback form still requires the installation of the wired system, which seems to have deterred more widespread use of this technique. Additionally, problems in understanding can only be diagnosed when a question is answered incorrectly as there is also no facility for students to initiate feedback - any problems in student understanding early in a period of lecturing will not be disclosed until the next period of questioning.

A different form of technology was introduced in reaction to the Feedback Classroom scheme. It sought to be more simple, inexpensive and robust and was introduced at RAF Cosford [TAPL]. Each student was provided with a wooden cube, the faces of which were painted different colours. The lecturer could then ask questions with up to six possible answers. "To indicate his choice of answer the student rests both

elbows on his desk with the cube cupped in both hands below his chin"; students could also initiate feedback as the "black face of the cube may be shown at any time the student feels the need during the lesson to indicate 'I do not understand' and the white may be used to indicate 'All clear now'" [ibid.]. This technique for obtaining answers to questions accords with the practical and economic considerations previously identified; however, the feedback is not completely confidential. This could lead to difficulties similar to those of 'show of hands' feedback. Additionally, students' indication of understanding difficulties whilst the discourse is in progress is not apposite. Students are required to disengage from any note taking, locate the black face and then go through the display procedure, thereby identifying themselves as having difficulties in comprehension (as the 'All clear now' function is necessarily preceded by a 'I do not understand'). It is also difficult to achieve much accuracy in the determination of the proportions of students indicating a particular answer.

There are no published analyses of the outcome of lectures using Cosford Cubes, and the technique has fallen into disuse despite its advantages of cheapness and robustness. It is likely that this is principally due to it having few advantages over a show of hands as a means of question response: it still requires students to identify themselves with opinions. This difficulty is more acute when students wish to inform the lecturer of comprehension difficulties during a lecture discourse.

3.3 Conclusions

Previous successful and unsuccessful educational technology innovations have been critically analysed. From these, and from wider studies, it is clear that an educational technology innovation should provide a solution on a specific weakness in a teaching situation, and be capable of introduction and use with the minimum of outlay of time or cost.

The most common formal teaching format of university education is the lecture. The clearest weakness of the lecturing modality is that it provides little opportunity for students' understanding to be appraised, and that students have little part to play in

the process. These weaknesses can be addressed by the introduction of a lecture technology to permit student feedback. This feedback should be available when the student feels it necessary to indicate difficulties, and permit the lecturer to pose questions to explore student conceptions.

The exploration of previous attempts to introduce lecture feedback serve to reinforce the conclusions drawn from the wider scale study of educational technology. The most successful of these, Audience Paced Feedback, is robust, simple, confidential, requires little effort to use from either the lecturer or students and has been shown to have a encouragingly positive effect on learning outcomes. However, this technology does require the installation of fixed infrastructure, which could have discouraged its more widespread introduction. Other feedback techniques have met with less success, primarily because they require greater input from students or lecturers.

The feedback introduced should, therefore, be as robust and inexpensive as possible. To obviate the need for fixed installations the feedback technology should be portable. The display of the students' opinions should be designed so that it can easily cope with large classes. A particular student's transmission of an opinion to the lecturer should be confidential to all lecture theatre participants; and all those present in a lecture should be able to initiate contact. The feedback technology should require little or no maintenance, to minimise the costs of delivery; the time and cost of the preparation of teaching materials to be used with the feedback must also be minimised.

Chapter 4   Overview of the feedback system

"might I point out to you that where it is not to all intents and purposes irrelevant, money is power, money is influence, money is effect"   Iain M. Banks, 'Excession'

The analysis of educational technology conducted in chapter 3 indicated that the most appropriate educational technologies with which to improve the lecture are those that are: robust, inexpensive, portable, capable of expansion to large class sizes, as confidential as possible to all lecture theatre participants and capable being used at both the students' and lecturer's behest. The technology should require little or no maintenance, and the preparation and delivery of teaching materials to be used with it should also require as small an outlay of time and cost as possible.

It was determined that these criteria can be met by providing each student with an inexpensive handset, similar in size to a television remote control. Upon these will be a number of buttons, which will be pressed by students to respond to multiple choice questions posed by the lecturer; they will also be used to allow the students to indicate that they are happy or unhappy with the speed at which the material is being presented during the lecturer's discourse. The data will be transmitted by radio to a central unit to remove the need for hardwired connections. Upon reception, the data will be aggregated and presented graphically to the lecturer for speedy assimilation, without identification of the transmitting handset. This graphical display will show the student's button presses as a percentage of the class to permit varying class sizes (the students 'register' their handset when they arrive at the lecture, with the total set as one hundred percent). The delivery of the feedback in the teaching situation will not be an onerous task as the lecturer can prepare multiple choice questions (for example, on additional overhead projections). The students will also be able to transmit data on the progress of their understanding with confidentiality and little difficulty.

The handsets must use as little power as possible to minimise any maintenance requirement (such as replacing or recharging batteries). This power requirement is

dependent upon the energy dissipated per transmission, and the transmission rate. In the first instance it will be assumed that each handset will be used at a maximum rate of once per second.

It was also decided that it would be extremely useful if the data transmitted were amenable to later analysis. To this end, each handset will also transmit an address which will be stored to hard disk with the data and the time at which it was received. This permits analysis of individual *handset's* action without the lecturer being able to identify an individual student with that handset.[1]

The system must be capable of providing every student in a class with a handset. The number of students that the system should cater for was initially fixed at two hundred and fifty six - for the twin reasons that this is 8 bits of binary communication and is a large class size in UK tertiary education. This number is not a final limit and represents only the interim target.

4.1 Educational overview

The feedback system will be able to be used by teachers who subscribe to any, all or none of the educational theories outlined in chapter 2. It provides the means to interact with students - and this interaction can take place without reference to the paradigm within which teaching is taking place.

As it is difficult (and inappropriate) to dictate the specific teaching style of the lecturer it was decided that the student-prompted feedback must be able to respond to something that can be varied by teachers irrespective of the teaching style. The feature that was identified as being common to all teaching methods is the speed at which material is being presented: if it is too fast for students to comprehend they may become confused, if it is too slow they may become bored. Therefore, it was decided in the first instance that this function should allow the students to indicate that the speed of delivery is too fast, about right or too slow.

---

[1] without specific action on the lecturer's part, such as numbering each handset visibly and recording which was used by particular students and referring to the stored data

The lecturer-prompted feedback could be used in many ways. It can be used to prompt for feedback on students' own opinion of their understanding ("Do you understand my arguments?") or it can be used to organise lectures ("Are you ready for me to continue?"), or in many other ways that may or may not be predicted. To allow the lecturer maximum flexibility in the nature of the questions which can be asked, it was decided that the lecturer-prompted feedback should be able to ask for multiple choice answers. In the first instance it was decided that the principal use of this function will be to pose questions that explore student understanding. This exploration could include asking for factual recall on the part of the students; however, as the majority of academics would agree that purely recalling facts is a comparatively minor part of a university education, the questioning should also attempt to encourage reflection on the part of the students. This could include the following categories of question (and this is not an exhaustive list!):

- identifying links between ideas and concepts between areas of study
- critically analysing the relative importance of a number of influences on a result
- transferring concepts from a known (taught) situation to an unknown one
- selecting an appropriate analytical technique with which to approach a problem presented to them

It is important to note that this function can be rapidly adapted by lecturers simply assigning new meanings to particular button presses (e.g. "Please press 'A' if you want more examples, 'B' if you do not, or 'C' if you have not understood this concept").

4.2 Economic overview

Universities have a much reduced resource base from which to purchase any new technology to improve each student's teaching, as discussed in chapter 1. It is therefore critical that the provision of each student-lecturer communication path be as cost effective as possible. It has been determined that this path should permit student prompted feedback on the speed of the lecture, and allow the lecturer to pose

multiple choice questions. This is a comparatively simple goal to achieve - and to attempt to exceed this will introduce unnecessary complication and possibly expense.

The system is also required to be robust. This indicates that the components used (particularly in the handsets) must be 'off the shelf' as these use technologies that are both cheap and proven. To minimise the costs of maintenance the handsets must also use the lowest power techniques available - a technical issue.

4.3 Technical overview

The handsets will operate on a duty cycle: ordinarily as inactive as possible, it will only transmit data when a button is pressed, with the time spent activated as short as possible. After transmission the button will be 'locked out' and unable to initiate another transmission until it is released, to prevent repeated transmissions and unnecessary power consumption from prolonged button presses.

The receiver can include complex and expensive electronics, if necessary, as there will only be one per system. It will be active continuously, but will have access to mains power; its power consumption is not constrained and it poses few technical difficulties.

4.3.1 Handset power consumption

The technology of choice when constructing low power inexpensive devices is CMOS, as it only dissipates significant energy whilst its logical state is being changed [HORO pp.154-156][TSIV]. The particular CMOS family that will be used (where discrete logic devices are required) is 74HCMOS, as this technology is reliable, inexpensive and the system will not need advanced technological capabilities [HORO pp.570]. The energy dissipation, E, of a CMOS device can be calculated by multiplying the time it is active, t, by the standard equation of CMOS power consumption [HCMO], giving:

$$E = \{f_iC_{PD}V^2 + \Sigma(f_oC_LV^2)\}t \qquad \text{Eq. 4.1}$$

where V is the supply voltage, C_{PD} is the input capacitance of the device, f_i the average (mean) frequency of input changes (between 0 and V), C_L the capacitative load of the device's output, f_o the average frequency of output state changes for this capacitative load.

There are two common techniques for the operation of digital circuits: synchronous and asynchronous. Synchronous circuits are designed so that each device operates simultaneously under the direction of a master clock. This technique ensures that there are no 'logic races' where a signal may be lost because it was present at the input of a device such that the device cannot read it[2] [HORO pp.515, 552]. However, synchronous circuits increase the frequency of input state changes as every input is read on each cycle of the clock, whether or not it has changed. The alternative technique, asynchronous operation, does not use a master clock. If this is used the designer has to be careful to ensure that the circuit does not exhibit logic races, but if the circuit does work correctly the factors ($f_i \times t$) and ($f_o \times t$) in equation 4.1 become constant - the number of input and output state changes becomes fixed and can be minimised. Therefore the handset was designed to operate asynchronously - without a master clock.

The energy dissipation of the handset has quadratic dependence on the supply voltage - clearly this voltage must be as low as possible. The specific logic family chosen, 74HCMOS, can operate from a supply ranging from 2 to 7 Volts [HCMO]. To guarantee supply fluctuations are within the specifications of the device - ensuring robustness - the handset will operate from a supply of 3 Volts where possible. Finally, the capacitance which the CMOS device is required to drive should be minimised (although in most cases any given CMOS device will be driving another CMOS device, which are designed have small input capacitance - in the order of pF [ibid.]).

---

[2] such as the 'removal time' - after a signal has been removed there is a finite delay before a new input can be read; or the 'set up time' which is the time required for the device to successfully change its outputs in response to its input

4.3.2 Handset power source

Using a photovoltaic power source would be ideal, as this would not require any maintenance during the handset's lifetime. The handsets will be operated in an artificially lit indoor environment where the light intensity can vary from perhaps 300 to a few thousand Lux (that is, around 1% of the outdoor illuminance where 1×10^5 Lux could be expected) [TAKA]. The most common light source in lecture theatres are fluorescent tubes.

There are three forms of commercially available silicon solar cells: monocrystalline, polycrystalline and amorphous [LUQU]. Amorphous cells have a comparatively low efficiency, with less than 6% of the incident radiation with outdoor illuminance being converted to usable power [LASN], compared with over 20% for monocrystalline cells [TREB] (with polycrystalline efficiencies in between). However, the absorption spectrum of the amorphous cells matches the emission spectrum of fluorescent lights more closely than the crystalline cells, resulting in this technology providing the most power under the conditions that the feedback system will operate [TAKA].

To obtain an up to date estimation of the order of magnitude of the power offered by commercially available amorphous silicon solar cells a number were obtained from a UK manufacturer [INTE]. Each of these units was 7cm long and 5cm wide, roughly the practicable area for energy generation on a handset, and consisted of an array of five individual 1cm wide cells connected in series. The cells were exposed to varying illuminances from a fluorescent tube in a darkroom (varied by changing the distance between the cell unit and a fluorescent tube, with the illuminance measured by a photometer). The maximum voltage (open circuit) and maximum current (short circuit) were measured. The maximum power output of the cells is somewhat less than this, figure 4.1; the typical amorphous silicon solar cell "fill factor" (the fraction of the product of the maximum voltage and current that can be obtained) is 0.6 [LASN].

Maximum power point

Short
circuit

Current

Voltage   Open circuit
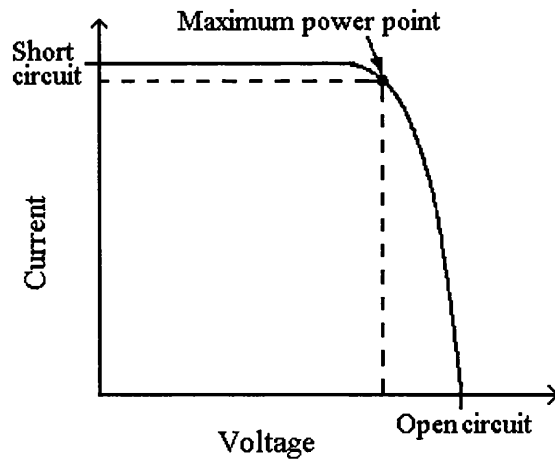
Figure 4.1 -- The output characteristics of a solar cell with a varying resistive load; the maximum available voltage is the open circuit voltage, the maximum available current is the short circuit current. The point of maximum power output is marked; the "fill factor" is the proportion of the product of the maximum voltage and current that can be realised at the maximum power point

The range of power outputs from these solar cell units is shown in figure 4.2. It is clear that photovoltaics under indoor fluorescent lighting conditions become a reasonable possibility as a source of power for the handset when its power consumption falls into the hundreds of microwatts (μW) range. As maximum transmission rate of the handset is once per second this requires the energy dissipation per transmission to be at most 200 microjoules (μJ).

As will emerge in appendix 8, the power dissipation is a factor of 6 greater than this. However, the design could be developed so as to achieve an energy requirement which would permit the use of photovoltaics. For our purposes, however, it was decided that the handsets would resort to batteries as a power source. Batteries that are sufficiently small and inexpensive to be used by a handset ("AA" size) are commercially available with an energy reservoir of some 15kJ if non-rechargeable (Alkaline), and 4kJ if rechargeable (NiCd) [DURA]. As recharging batteries is a maintenance requirement it was decided that non-rechargeable batteries would be used.
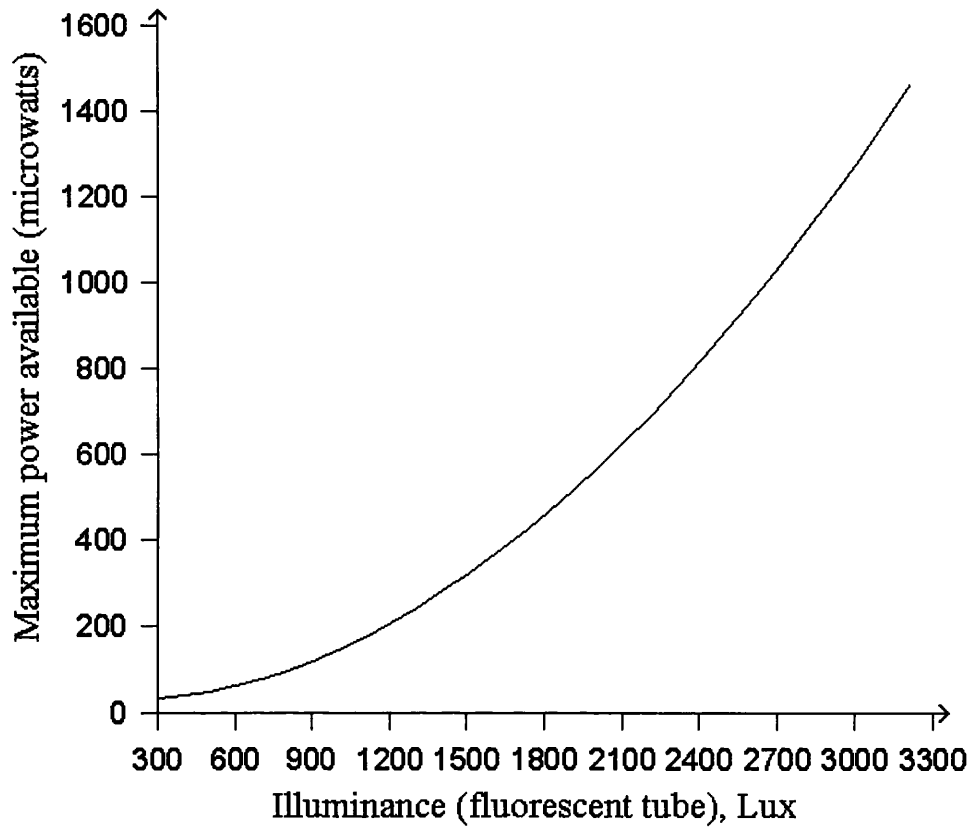
Figure 4.2 -- The experimental results of an order of magnitude estimation of the power supply available from an amorphous silicon photovoltaic unit (consisting of five 7cm individual cells connected in series) under indoor illumination levels from a fluorescent tube. This represents the open circuit voltage multiplied by the short circuit current multiplied by the fill factor (0.6)

Chapter 5 Networks

"They've eradicated unemployment now. Mind you, they've raised the school leaving age to sixty-nine" Alexei Sayle

The system should be designed to consume the minimum possible energy per data bit, with the further caveat that the data should have an acceptably low probability of being received in error. This chapter examines avenues to approach this on a network scale; the following chapter explores the mitigation of the effects of random noise.

Data transmission channels have a fundamental property: 'capacity', reflecting the maximum possible amount of data that can be transmitted per unit time.[1] The proportion of the channel capacity that is being successfully used is channel 'utilisation'; and channel 'traffic' is the total amount of data that being transmitted whether successfully received *or not*.

5.1 Clock capture

Received signals are just meaningless voltage levels without the recovery of the clock with which to ascertain the data bit duration. A channel separate to the data could be devoted to the transmission of this information - an unattractive option in a power-limited system. Therefore this information is sent on the same channel as the data.

5.1.1 Synchronous capture

The data could modulate the carrier directly, with the clock recovered from data transitions between high and low. However, the receiver clock could drift from

---

[1] Shannon's capacity expression indicates the absolute limit of a channel's capacity - with bandwidth W and a signal to noise ratio S/N this limit is $W \log_2\{1 + S/N\}$ [SHAN]. To reach this limit requires formidably complex communication systems; in practice channels have capacities rather less than this [SCHW pp.631-636]

synchronisation if there is a sufficiently long series of identical bits. Synchronous clock recovery is the concurrent provision of the clock with the data; there are a variety of means effecting this [BYLA pp.190-218, 241-246]. One method is to use Manchester coding, figure 5.1, where two bits of opposite polarity represent a single data bit, with the central transition used to recapture the data clock [SCHW pp.355-357]. Simple Manchester coding is where a data high is '10' and a data low '01' (or vice versa). Differential Manchester coding is where a data 'high' retains the same signal level as the last half of the previous bit then inverts it; a data 'low' has transitions at the beginning and middle of the data bit. This technique can tolerate indefinite data streams, as the clock is continually refreshed, but doubles the number of bits transmitted. More advanced and involved codes can reduce the number of redundant bits it is necessary to transmit [BYLA pp.340-344]. There is, however, another solution: asynchronous clock capture.

| Data stream | 1 | 1 | 0 | 1 | 0 |

Simple Manchester coding

Differential Manchester coding
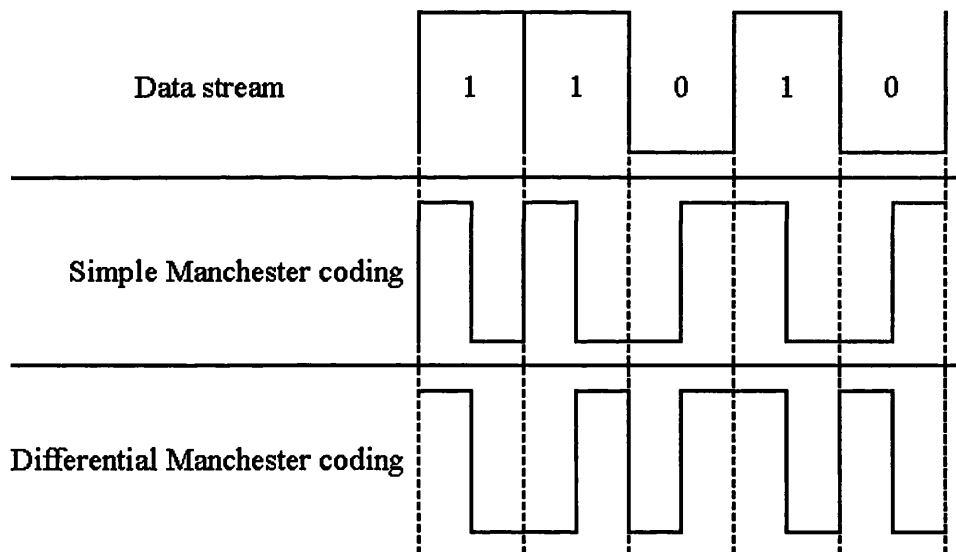
Figure 5.1 -- Simple and differential Manchester coding

5.1.2 Asynchronous capture

Asynchronous clock recovery is the provision of a 'mark' bit to synchronise the receiver clock with the transmitting clock. The system presumes that these clocks will not drift sufficiently far from alignment to cause error during subsequent data transmission. This requires precise clocks - but inexpensive crystal oscillators with drifts of parts per million are readily available. This clock capture technique

constrains the number of bits it is possible to send successfully between mark bits, but only necessitates the transmission of a single bit to recover the clock.

5.1.3 System choice

The feedback system is required consume as little energy as possible and transmits small amounts of data (eleven bits per transmission). It was decided that asynchronous clock capture, minimising transmission duration and hence power consumption, would be used.

5.2 Random networks

The feedback system consists of many handsets transmitting to a single central receiver. If handsets simply transmit when there is data (access to the communication channel is random) there is a chance that two handset transmission will overlap (a 'collision'), causing data loss. However, this approach has the merit that is the simplest possible, requiring a minimal number of components and hence minimising power dissipation.

Calculation, conducted in appendix 1 (see also [KLEI]), indicates that the probability of collision for a random access handset transmission of length τ, in a system of N handsets is:

$$P_C = 1 - e^{-2kN\tau} \qquad\qquad \text{Eq. 5.1}$$

where k is the number of transmissions per handset per unit time.

5.3 Organised networks

There is an obvious solution to the collision problem: organising handsets' access to the channel to ensure successful transmissions. There are a variety of techniques that could be used in the system, which are outlined below; throughout this section

quoted figures are for a classroom situation with a 256 handset system, each transmitting once per second.

5.3.1 Time division multiple access

Time Division Multiple Access (TDMA) is the allocation of time slots for the transmission of data from a handset - access to the transmission medium is timetabled [COMS]. A period, 'guard time', is allocated on either side of each slot to ensure that no two transmissions overlap. The total number of slots plus guard times for the handsets is a 'frame'.

TDMA is particularly efficient at ensuring that the transmission channel is utilised effectively when there is very high demand; each handset transmits in every one of its allocated slots - the total amount of data transmitted is very close to the channel capacity. The timing data for access can be provided by external clock signals or by internal clocks in the handsets periodically synchronised such that any clock drift is less than half of the guard slot. The former option is not alluring: for a robust link the timing communication channel has to be at radio frequencies (as infrared could be blocked for lengthy periods by bags, books, bodies etc.). The receiver in the handset would need to be on continuously - a handset power requirement of 165 milliwatts (mW) using a commercially available low power receiver [RF99r] with the local oscillator being provided by [RF99t]. The alternative internal clock would dissipate in the order of $35\mu W$.[2] Its periodic synchronisation also requires a robust channel (as an error in the synchronisation signal would be highly likely to cause collisions) and would have to take place frequently enough to negate the effects of clock drift. With a frame of 1s the 256 handsets can be separated by 3.8ms; the handsets' clocks could drift toward each other: the guard time is 1.9ms. This necessitates clock synchronisation every 300s for a crystal with a drift of 5 parts per

---

[2] the clock is a low power crystal oscillator module and two counters. Crystal oscillator power consumption is calculated from [HA72 pp.4, 9]. The counters are standard [HC404] with a load capacitance of 20pF and input capacitance of 3.5pF. The oscillator operates at its 10kHz minimum; the first counter input frequency is 10kHz, output frequency 5kHz; the second counter input frequency is 5kHz, output frequency 1Hz. The use of standard logic elements is assumed, with mean clocking frequency of 250Hz

million,[3] raising the total average internal-clock TDMA power requirement to 46μW.[4]

5.3.2 Token passing

Token passing is a channel access technique in which there is distributed control: each station is allocated a number or position on the network and permission to transmit is passed sequentially between them by the use of a 'token'. If a station has no data to transmit the token is immediately passed on the next station. The token can be passed directly from handset to handset, or via a mediating central unit.

The use of token passing is particularly useful for networks where the probability that every station will have data to transmit when the token is passed to it is close to, but not quite, unity. It is akin to TDMA as a form of timetabling, but with the means to instantaneously match the supply of channel access to handset's demands. As the time at which the token will be passed to a handset will vary if other handsets have not had data to transmit, each handset needs to have an continuously active receiver to accept it. This receiver needs to be reliable, as no data is transmitted from the handset until the token is passed to it, resulting in a power requirement of 165mW as before. The handsets could be 'polled' for data periodically (perhaps once per second) which would decrease this power requirement as an internal clock could be used to activate the handset's receiver only during this polling. The receiver has to be active for at least 20ms for reliable reception [RF99t][RF99r]. Neglecting the time needed for the token passing itself, this results in a power consumption of 3.3mW (with the assumption that the handset's internal clock is synchronised during the polling procedure).

---

[3] this drift is that quoted for [HA72] operating at 10kHz. The eighty second reduction from the theoretical maximum time between clock synchronisation signals, 380s, is to *ensure* that no overlaps occur

[4] to ensure successful reception of the synchronisation signal the robust receiver - that used to receive the external clock - needs to be active for 20ms to stabilise; an energy consumption of 3.3mJ. Averaged over 300s this represents a power requirement of 11μW

5.3.3 Collision sense

Collision sense network access is purely demand based; when any handset has data to transmit it monitors the channel for collisions - either avoiding or detecting them. Collision avoidance handsets observe the transmission channel and transmit when it is unused. With collision detection the handset transmits immediately, and then waits for confirmation from the receiver that transmission was received correctly. If it was not, the handset retransmits some time later (with delays before retransmission different in each handset to avoid the previous collision).

These systems differ in the expected channel utilisation. With collision avoidance schemes, it is presumed that that the system might be in use at any given moment. However, it is further assumed that it is *unlikely* that two handsets will wish to transmit during the pre-existing handset transmission - as if this occurs, a collision is guaranteed after the end of this transmission. In a collision detection system a lower channel utilisation is expected - the channel is anticipated to be unused - and the handsets transmit in the expectation of success. Further, there is the presumption that the probability of collision is low enough for retransmission to be likely to succeed. In fact, collision detection systems become unstable above a certain channel utilisation; retransmissions collide with other transmissions or retransmissions and recur, increasing the amount of traffic on the channel whilst decreasing its utilisation. Calculations under certain assumptions on one representative network (the Aloha system, appendix 2) show that the maximum channel utilisation before instability is about 18%. There are techniques to increase this limit, such as the abortion of transmissions when multiple collisions are detected, but a channel cannot be utilised to capacity with collision detection techniques [LAMS].

Collision sense is the most flexible form of network access - the number of active handsets is not predetermined and can change instantaneously (provided the upper bound on the amount of data transmitted is not breached). Each handset has to

include a receiver (as it must be able to monitor the channel/be reliably informed of collisions) - a power requirement of at least 3.3mJ per data transmission.[5]

5.3.4 Frequency division multiple access

Frequency Division Multiple Access (FDMA) side-steps the question of access by providing a channel for each handset (or small number of handsets). The channels operate at spaced frequencies reducing, or even obviating, the need for the organisation of network access.

5.3.5 System choice

FDMA is an elegant solution to the network access problem; however, the system is likely to use a relatively low frequency for transmission (less than one gigaHertz) to avoid any need for high-tech, expensive devices in the handsets. These frequencies are both heavily used and regulated. FDMA also necessitates a far larger bandwidth than other techniques potentially causing problems with noise from other high power radio-frequency devices at some frequencies. It also limits the achievable economies of scale, as each handset has to have a different transmitter, or at the very least different crystals and external components for their transmitters. For these reasons the option of FDMA was ruled out.

The handsets with organised network access dissipate more power than those that endure collisions with random access. For this reason, no access technique has been used in the current feedback system - the primary aim must be to reach the power consumption criterion before any more advanced techniques are used. In addition, there exists a more attractive option permitting simultaneous access for the handsets, code division multiple access, which also provides a degree of protection against random noise, section 6.2.

---

[5] again the receiver must be active for 20ms to ensure reliable operation [RF99r][RF99t]. This results in the consumption *at least* 3.3mJ per transmission (as if there are collision problems the handset receiver could need to be powered for longer)

5.4 Spread spectrum

Spread spectrum, figure 5.2, is "a technique whereby an already modulated signal is modulated a *second* time in such a way as to produce a waveform which interferes in a barely noticeable way with any other signals operating in the same frequency band" [TAUB p.720]. Users can access the channel *simultaneously* without the extreme degradation of the probability of error that would be expected with conventional modulation. Code Division Multiple Access (CDMA) is the name given to the application of spread spectrum in single-channel, multiple-user environment.
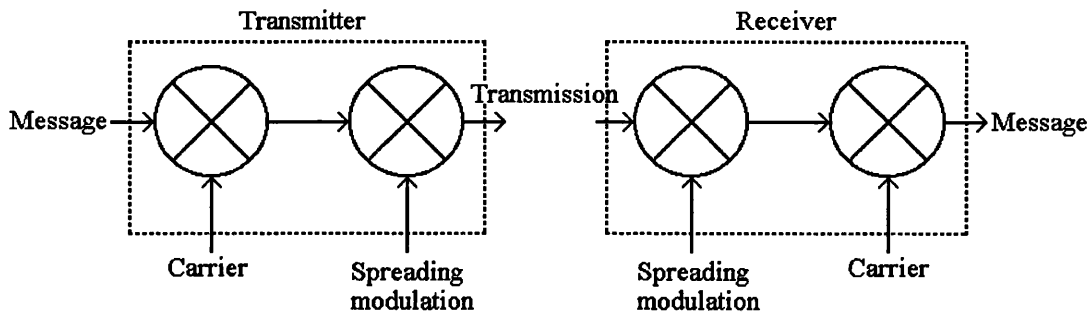
Figure 5.2 -- A general block diagram of a spread spectrum communication system

5.4.1 Pseudonoise

The most common form of spreading modulation is 'pseudorandom' codes [COOP pp.274-275]. These binary codes have the instantaneous appearance of noise, but in fact have defined properties providing multiple access capabilities. Chirp spread spectrum, which uses a spreading modulation to sweep each transmitted bit through transmission frequencies, does not provide attractive multiple access performance and is principally used for its noise rejection properties [DIXO pp.44-47].

The multiple access properties of a CDMA system are the result of code used to spread the modulated waveform. One class of codes, 'maximal length', have desirable properties [DIXO pp.58-64]. These codes are maximal as they are the longest codes that can be generated by a bank of shift registers, figure 5.3; the output code of a bank of Q shift registers repeats after $(2^Q - 1)$ bits.
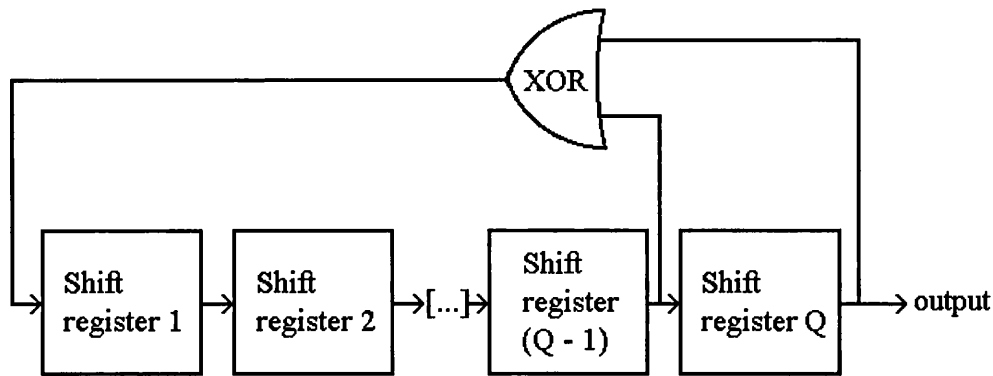
Figure 5.3 -- Pseudo-noise generation using a bank of Q shift registers

Maximal length codes exhibit low correlation between bit-shifted versions of themselves, low 'autocorrelation'. This property can be demonstrated for a three stage shift register sequence, with the reference sequence 1110010:

Shift	Sequence	Agreements (A)	Disagreements (D)	(A - D)
1	0111001	3	4	-1
2	1011100	3	4	-1
3	0101110	3	4	-1
4	0010111	3	4	-1
5	1001011	3	4	-1
6	1100101	3	4	-1
0	1110010	7	0	7

Table 5.1 -- The autocorrelation properties of a three stage shift register maximal length code with respect to a reference sequence '1110010'

This autocorrelation property is common to all maximal length code sequences and is extremely useful for some systems. However, in a multiple user environment *cross-correlation* between users is the critical variable. The maximum level of cross-correlation between maximal length codes has not been amenable to analysis. Potentially, it could rise to levels causing the receiver to decode in error - although some results are promising [TIRK].

There is a solution: Gold codes, whilst having sub-optimal autocorrelation, *do* have bounded levels of cross-correlation [GOLD][SARW]. These codes are constructed from maximal length codes, figure 5.4, to specified criteria to ensure that the cross-correlation is optimised.[6] This provides multiple access capability: the correlation of a correct code with received signal's embedded code will be very much higher than any other correlation. This permits the both the identification of the handset from which transmission originated and the data sent. The precise cross- and autocorrelation properties of the Gold codes to be employed in a spread spectrum feedback system can be calculated [FANF]. These codes modulate the data by various techniques, including: 'frequency hopping', 'time hopping' and 'direct sequence'.
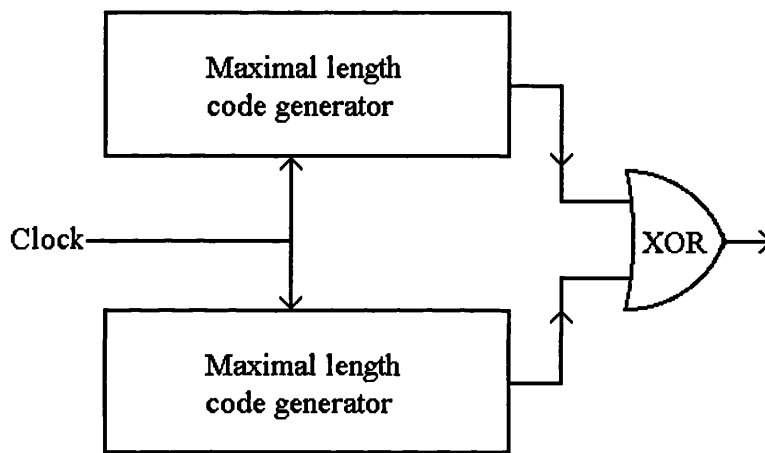
Figure 5.4 -- The generation of Gold codes

5.4.2 Frequency hopping

Frequency hopping CDMA is the use of sections of the pseudonoise code to synthesise discrete frequencies, which are mixed with the data-carrying carrier - essentially, the carrier jumps about in frequency as directed by the code, figure 5.5. The receiver reverses the process to recover the baseband modulation. Frequency hopping has the advantage that it can be programmed to avoid noisy portions of the spectrum. However, it requires frequency synthesising circuitry in the handset [DIXO pp.126-136]. In addition, there is a minimum bound on the probability of

---

[6] an algorithm with which to choose Gold codes with good cross-correlation performance is given in [DIXO Appendix 7]

error in the reception of a bit - if narrowband noise interferes with a particular frequency all the information sent on that frequency will be lost; only the introduction of error correcting/detecting codes (section 6.3) can reduce this probability further.
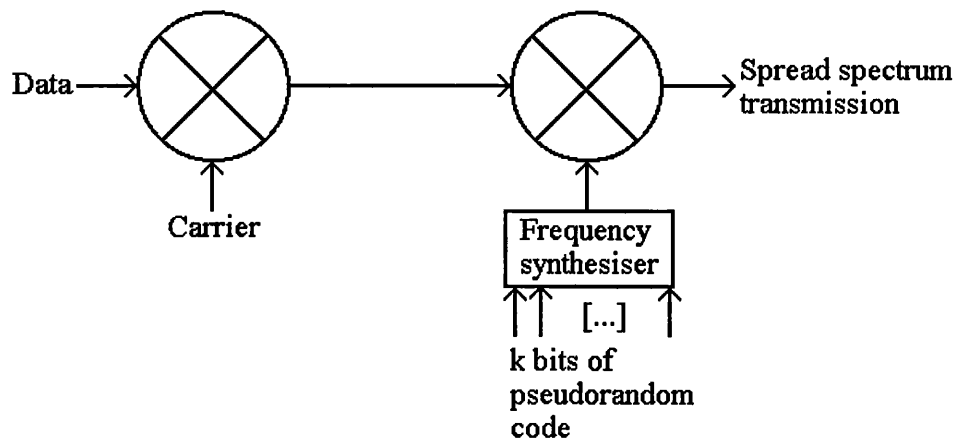
Figure 5.5 -- The generation of a frequency hopping spread spectrum signal

5.4.3 Time hopping

Instead of using the pseudorandom code to hop about in frequency it can be used to hop about in time [DIXO pp.42-44]. If time is divided into frames, and these frames are further divided into j slots, each receiver can hop between slots as directed by $(\log_2 j)$ bits of the pseudorandom code. This spread spectrum method is simpler to implement in the transmitter than frequency hopping, but has the critical disadvantage that a universal clock is needed for all of the system. This generates a power demand equal to that of internal-clock TDMA (section 5.3.1).

5.4.4 Direct sequence

The most attractive form of spread spectrum for the feedback system is Direct Sequence CDMA (DS-CDMA), which applies the spreading modulation directly to the data, figure 5.6. The clocking frequency of the pseudorandom code is very much faster than that of the data; additionally it is common for the pseudorandom code and

44

data to take the values ±1. The pseudorandom code is modulo-2 added to the data,[7] and the result modulates the carrier. Upon reception this process can be reversed to recover the data. This technique has the advantage that it is simple to implement in a handset, requiring only the addition the pseudorandom code generating circuitry and a modulo-2 adder; it also has excellent narrowband noise rejection properties (section 6.2).
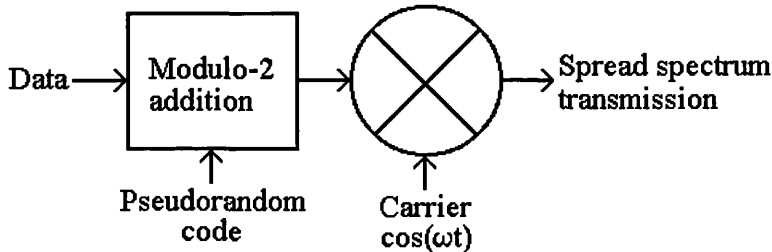
Figure 5.6 -- The generation of a direct sequence spread spectrum signal

DS-CDMA had been regarded as a weak candidate for spread spectrum systems, as it appeared to suffer from the 'near-far problem'. This problem occurs when two handsets, at different distances from the receiver, transmit simultaneously. Whilst the cross-correlation level between handset codes is low, it is not zero. When this low correlation is mixed with the high received power from a nearby handset the resultant voltage may exceed the voltage generated from the correct code mixed with the lower power received from the farther handset. This difficulty seemed to circumscribe the use of DS-CDMA to those areas where the remote transmitters are received at approximately equal powers (such in terrestrial-satellite communications).

However, DS-CDMA does become viable if the handsets vary the power with which they transmit. The receiver periodically polls the active transmitters and ascertains the power received from each (and hence their distance). The receiver then informs each transmitter of its desired transmission power level (such that all are received at

---

[7] modulo-2 addition has the truth table for two signals which can have states A or B:

	A	B
A	A	B
B	B	A

the level of the weakest). This procedure has the advantage that all but one of the transmitters will save power by transmitting at a sub-maximum level; but has the critical disadvantage that every transmitter is also required to have a receiver.

Multiuser DS-CDMA obviates the need for this receiver by treating the received signals from all the transmitters *en masse* [VERDa]. An optimal multiuser detector knows or can acquire: the codes applied to all users' data, the timing of all users' signals (both bit timing and carrier phase) and the received amplitudes of all of the users interfering with the desired signal. This information is then used to regenerate the desired signal *without* the near-far problem. However, it can be shown that the complexity of such an optimal receiver rises as two to the power of channel utilisation[8] [VERDb]. This rapidly generates unworkable complexity in the receiver!

Recent research attention has therefore turned to realisable, sub-optimal multiuser detectors. One such approach is successive interference cancellation [BOTT][PATE]. The received signal is correlated with the users' codes, the strongest signal present is then regenerated and its data is decoded. This regenerated signal is removed from the incoming signal. That resultant signal is then again presented to the bank of correlators and the process repeated, figure 5.7. This successive cancellation of the interfering users' signals permits the reception of weak signals, but "the processor performing the cancellation must perform all the cancellations while maintaining the necessary data rate" [ALIF]. This might require the use of buffers for the incoming baseband signal if the required operations could not be conducted at sufficient speed. With our system having a low data rate this approach is promising.

---

[8] the number of users times number of bits transmitted per user per unit time

Signal →

Carrier
cos(ωt)

a_1(t)

$\int_0^T dt$

a_2(t)

$\int_0^T dt$

[...]

a_N(t)

$\int_0^T dt$

Select
largest
signal

strongest
→ signal
demodulated

Regenerated signal at received amplitude
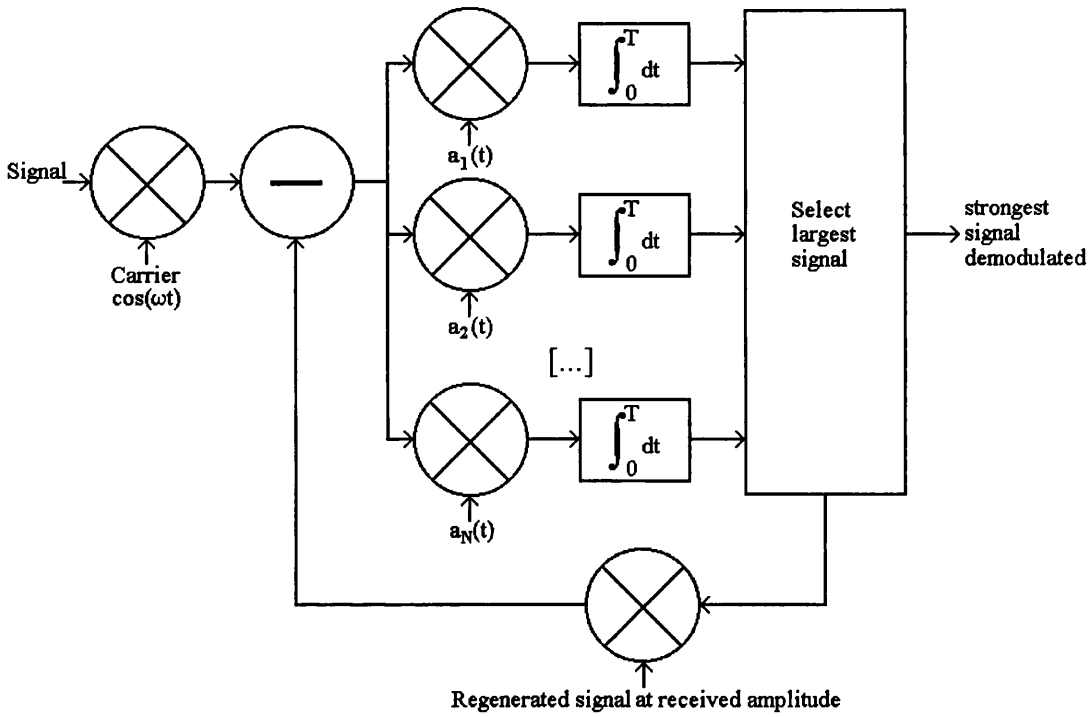
Figure 5.7 -- Block diagram of successive interference cancelling multiuser detector

5.4.5 CDMA system development

Throughout the preceding discussion on the forms of spread spectrum analysis of the difficulties arising from the coarse synchronisation of the user's code with the incoming stream (to within a single pseudonoise bit) and the subsequent tracking of the local pseudonoise (to within fractions of a pseudonoise bit) has been neglected as these problems are well documented in the literature with a variety of solutions [COOP pp.345-376][DIXO pp.214-260][GOLO][TAUB pp.738-745]. Additionally, as the utilisation of the channel approaches its capacity, spread spectrum system's probability of bit error *does* degrade, in a known manner [VITEa]. However, the total amount of data sent in the feedback system is very small so it is likely that channel utilisation will not rise to degrading levels.

5.5 Conclusions

The chief design principle of the feedback system is that it should be as economical as possible; both in terms of immediate capital cost and minimising subsequent

maintenance. The maintenance criterion requires the handsets to consume as little power as possible, to mitigate any need to replace or recharge power supplies (with the ultimate goal of using a photovoltaic supply).

The clock with which to read the transmitted data bits will be recovered using asynchronous techniques to minimise the number of bits it is necessary to transmit.

The feedback system is statistical in nature. It is intended to permit the majority view of the students to be apprehended by the lecturer - the correct reception of every individual transmission is not critical. In light of this, and as organised network access techniques increase the energy consumption per transmission, it was decided that handsets' access to the transmission channel would be random. The probability of error for the feedback system can be calculated using equation 5.1. A relatively high probability of collision can be endured: the initial tolerance was set at 5%. The design aim is to have a system with 256 handsets; it was further determined that the maximum rate of data input from the students would be one press per second (as could occur when students are answering a particularly simple question). Under these criteria each handset transmission in a simple random access system, by substitution into equation 5.1, should last less than $100\mu s$.

Spread spectrum techniques to remove the probability of collision for random access channels are alluring, particularly DS-CDMA as the complexity in this scheme is concentrated exclusively in the (comparatively power-rich) lecturer's receiver. In principle, with multiuser techniques, it should give superior performance to the random access approach. However, for the initial development of the feedback system spread spectrum will not be utilised to minimise the handset power requirement - and robust, inexpensive components with which to effect these techniques are not yet available.

Chapter 6   Noise

"Education [...] has produced a vast population able to read but unable to distinguish what is worth reading"   *G.M. Trevelyan*

Any transmitted signal, irrespective of the nature of the network over which it is transmitted, can be received in error because of noise in the transmission channel. Wideband noise can arise from a variety of sources, such as from the random motions of charge carriers within the receiver (thermal noise), or from external sources, such as man-made narrowband noise [ZEIM pp.747-754]. This chapter considers the variation of the probability of error due to noise. It is possible to both detect and (with the deployment of more sophisticated approaches) correct errors by the addition of extra code bits to the data transmission; simple forms of these codes are described.

6.1 Wideband noise

Wideband noise is commonly modelled as being 'Additive White Gaussian Noise' (AWGN) - that is the noise is simply added to the signal, has a constant power spectral density (hence "white"),[1] and its instantaneous magnitude (which can be negative) varies according to Gaussian statistics. At any given frequency the noise amplitude or phase, η, has the Gaussian probability density function $p(\eta)$

$$p(\eta) = \frac{e^{-\eta^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \qquad\qquad \text{Eq. 6.1}$$

with a mean, m, of zero and standard deviation σ. This is a justifiable and reasonably accurate representation of noise as a linear ensemble of very many individual random

---

[1] this is evidently not possible over all of the electromagnetic spectrum (as this requires infinite power) but can be shown to be a reasonable approximation for the effects of thermal noise over frequencies up to 10^{13} Hz [PROA p.190-191]

processes, whether Gaussian or not,[2] can be closely approximated as a single Gaussian process - the well known statistical Central Limit theorem [ZEIM pp.312-315]. The white noise is the result of the superposition of the Gaussian probability density functions over all frequencies; hence when this white noise is filtered the resultant spectrum is also Gaussian, with a 'noise-equivalent bandwidth' (that of an ideal rectangular bandpass filter with the characteristics of unfiltered noise within it, and zero elsewhere)[3] permitting the calculation of the total noise power [PROA pp.192-204][GRAY]. As the AWGN was formed of many other Gaussian spectra, this bandlimited noise can also be represented by equation 6.1 [ZEIM pp.362-370].

6.1.1 Binary symmetric channel

A binary channel is where two signals are transmitted representing a 'high' or 'low'; it is symmetric if the *a priori* probability of either a high or a low being transmitted is equal. These signals modulate the carrier with magnitudes of $\pm\alpha$, symmetric about some value μ (this modulation can be phase, frequency or amplitude). The received signals will have an additional noise component. The concept of 'signal space' permits the depiction of the probability that received signals will take a given value [SCHW pp.571-572]. The probability density function is plotted graphically upon axes, ϕ_i, representing modulation used (for example, if phase modulation is used the axis ϕ could be $\cos\{\theta\}$). A binary symmetric channel with received noisy signals plotted on a general signal axis ϕ is shown in figure 6.1. The receiver must decide, given a received signal level, which signal was *transmitted*: a decision threshold is set. The threshold, μ_T, for figure 6.1 is midway between the two signals as the probability of a high or low being transmitted is equal (and the time averaged noise does not bias the signals). When signals have differing *a priori* likelihoods of being high or low the threshold moves reflecting the resultant change in the *a posteriori* probability.[4]

---

[2] as long as they have a finite mean, non-zero standard deviation and each provides a small contribution to the total - such as the random motions of charge carriers causing thermal noise
[3] this bandwidth is usually provided by filter manufacturers
[4] the generalised mathematical derivation of the decision threshold as the *a priori* probabilities change is given in [STRE p.673-678]

Probability

−α Decision α φ
threshold, μ_T
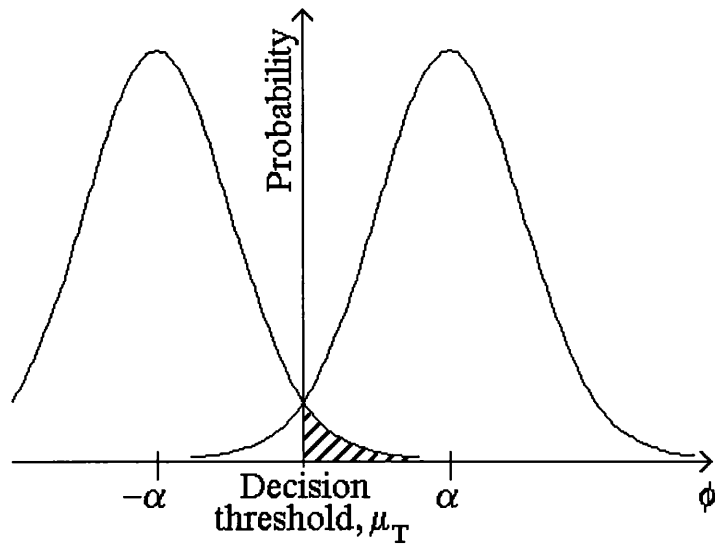
Figure 6.1 -- The probability density function of two received signals, with some modulation φ (phase, frequency or amplitude) with magnitude ±α about some level μ. The noise in the channel is Gaussian. If the receiver receives a signal to the right of the decision threshold (falling at $\mu_T = \mu$) it determines a high was sent. However, if a low had been sent this would be in error; the region of the probability density function where this occurs is shown shaded

The power necessary to generate a signal is proportional to the magnitude of the modulation squared.[5] An error occurs when the noise causes a signal to fall on the incorrect side of the threshold and be misapprehended, the shaded region for a low signal in figure 6.1. It should be noted that if the signal magnitudes were 3α and 5α the probability of error (the region of overlap) would be unaffected, but the total power needed for transmission would be much increased. This demonstrates a more general result: the power required for the transmission of any signal set is minimised when its centre of gravity in signal space lies at the origin.

The erroneous reception of a single transmitted 'low' is the shaded area of the Gaussian curve. This curve has a mean value of α, with variance σ (equation 6.1); therefore the probability of a transmitted low falling on the high side of the decision threshold, μ_T, is:

---

[5] since $P = V^2/R$

$$P_{e0} = \int_{\mu_T}^{\infty} \left(\frac{e^{-(\phi + \alpha)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \right) d\phi \qquad \text{Eq. 6.2}$$

With a corresponding probability of 'false dismissal' for a high falling on the low side of the decision threshold. However, the *net* probability of error is conditional upon the signal sent. With the probability of a low being transmitted P_0 and the probability of a high being P_1 the total probability of error for the channel is:

$$P_e = P_0 P_{e0} + P_1 P_{e1} = \tfrac{1}{2}(P_{e0} + P_{e1}) = P_{e0} \qquad \text{Eq. 6.3}$$

since $P_0 = P_1 = \tfrac{1}{2}$ and $P_{e0} = P_{e1}$ from symmetry. However, equation 6.2 is not evaluable in closed form: one must resort to numerical techniques. This is commonly conducted using $Q(\phi)$, a standard function[6] where

$$Q(\phi) = \frac{1}{\sqrt{2\pi}} \int_{\Psi}^{\infty} \left(e^{-z^2/2} \right) dz \qquad \text{Eq. 6.4}$$

for a generalised decision threshold Ψ and with a change of variable such that

$$z = \frac{\phi - m}{\sigma} \qquad \text{Eq. 6.5}$$

where m is the mean and σ the standard deviation of the Gaussian curve (i.e. the signal level) as before. Solution of equation 6.2 using equations 6.4 and 6.5 permits comparison between the error performance of various modulation techniques.

---

[6] further elucidated in [FRER]. Another standard function used to evaluate equation 6.2, the error function erf(ϕ), is defined in [SCHW p. 417]; the relation between the two is demonstrated (with tabular values of the Q-function) in [STRE p.746-748]

6.1.2 Error performance

There are two basic types of signal demodulation - coherent and non-coherent [HAMB pp.367-381]. In coherent reception the receiver synchronises its local carrier to the received carrier, thereby recovering phase information. In non-coherent demodulation the receiver detects the envelope of the signal by integrating it over time, obviating the need for phase synchronicity. As would be expected, non-coherent demodulation has a poorer error performance as the receiver disregards any phase information regarding the noise.[7] This chapter will concentrate on optimum, coherent demodulation when presenting noise mitigation techniques.

The most basic form of modulation is on-off keying where the presence of a signal indicates a high and its absence a low. This is a particular form of an amplitude shift keyed modulation, but is a general representation for this calculation as changing the energy levels such that the signals have an equally changed amplitude merely results in a translation of the signal set in signal space, not affecting the probability of error. This modulation has signals:

$$s_{LOW} = 0 \qquad\qquad s_{HIGH} = A\sin(\omega t) \qquad\qquad \text{Eq. 6.6}$$

Frequency shift keying is the modulation of the carrier frequency

$$s_{LOW} = A\sin(\omega_1 t) \qquad s_{HIGH} = A\sin(\omega_2 t) \qquad\qquad \text{Eq. 6.7}$$

and phase shift keying uses the phase of the carrier to transmit the information

$$s_{LOW} = A\sin(\omega t) \qquad s_{HIGH} = -A\sin(\omega t) \qquad\qquad \text{Eq. 6.8}$$

When the probability of error is calculated for the optimum coherent reception of these modulations with AWGN, phase shift keying provides the best performance, figure 6.2 [SCHW pp.587-592].

---

[7] to achieve the error performance of a coherent system, non-coherent demodulation requires just over a 3dB increase in the signal to noise ratio [HAMB p.376]
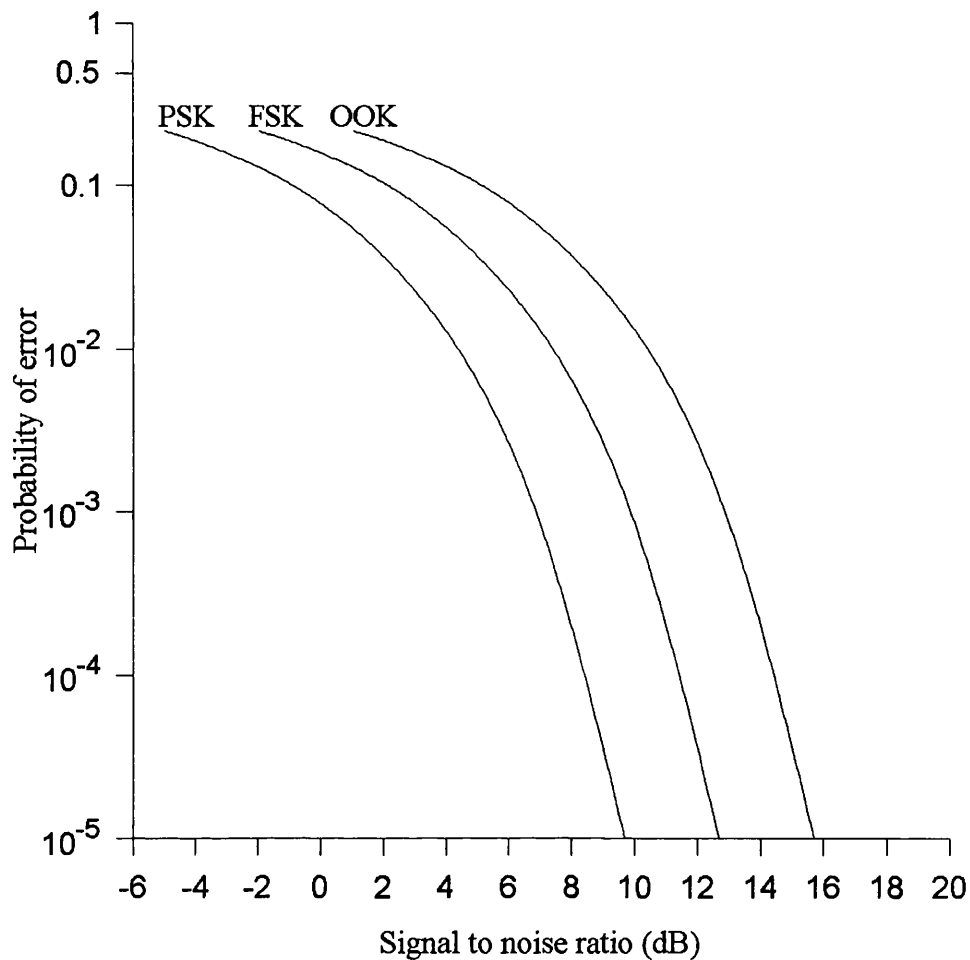
Figure 6.2 -- Comparative error performance of On-Off, Frequency Shift and Phase Shift Keyed modulations (OOK, FSK and PSK respectively) over a range of transmitted bit energies; as this energy tends to zero error probability tends to 0.5 (the a priori likelihood of a high/low being sent)

The unsurprising general conclusion of figure 6.2 - that increasing the signal to noise ratio improves error performance - reflects the increasing separation of the signals in signal space, figure 6.1, decreasing the size of the region in which overlap can occur. In practice there is a finite energy budget which limits this separation.

6.1.3 M-ary modulation

There is an alternative approach: the simultaneous transmission of several bits in the same signal. The resulting M symbols, formed from j bits of information, number 2^j. Each symbol has a transmission energy of j times the bit energy, figure 6.3; this is known as 'M-ary' modulation. For example, if two bits (with modulation magnitude α) are combined and transmitted using phase shift modulation the signal set, 'Quadrature Phase Shift Keyed' (QPSK), would be

$$s_1 = \alpha\cos(\omega t) + \alpha\sin(\omega t) \qquad \text{representing '11'}$$

$$s_2 = \alpha\cos(\omega t) - \alpha\sin(\omega t) \qquad \text{representing '10'}$$

$$s_3 = -\alpha\cos(\omega t) + \alpha\sin(\omega t) \qquad \text{representing '01'}$$

$$s_4 = -\alpha\cos(\omega t) - \alpha\sin(\omega t) \qquad \text{representing '00'} \qquad \text{Eqs. 6.9}$$
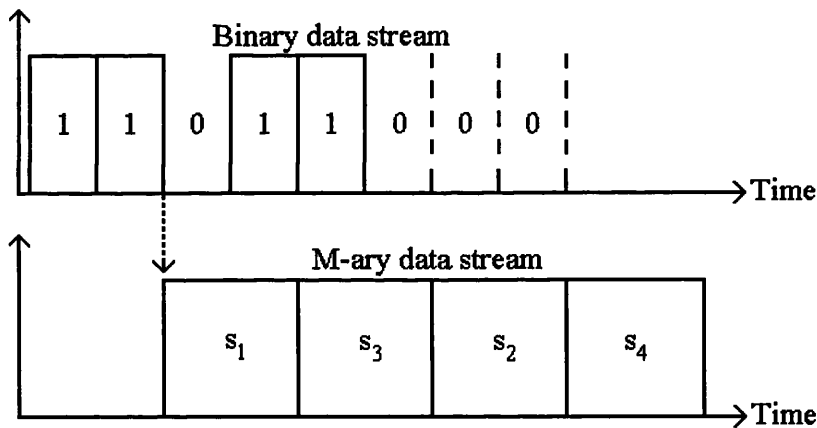
Figure 6.3 -- The conversion of a binary data stream to a quaternary data stream

that is, organised so that the most likely receiver symbol misapprehensions result in a single bit error, figure 6.4. This signal constellation can be viewed a two simultaneous, independent and non-interfering orthonormal binary phase shift keyed signals since the functions $\cos(\omega t)$ and $\sin(\omega t)$ are orthogonal and can be normalised by the multiplication of each by a factor of $\sqrt{(2/E_S)}$, where E_S is the symbol energy. The orthogonality can perhaps be best visualised by taking the noise to be the projection of a rotating phasor upon two axes in a two-dimensional plane. As the phasor rotates in either direction the projection on the horizontal axis, $\sin(\omega t)$, does

not interfere with the projection on the vertical axis, cos(ωt). The first half of each symbol can be thought of as permitting a demodulator decision upon the cos(ωt) component and the second half discerning the sin(ωt) component. As the energy per *bit* remains constant (the duration of each symbol is twice that of each bit) the error performance of this quaternary phase shift keyed set is the same as that provided by a simple phase shift keyed modulation - the 'PSK' curve of figure 6.2 [ZEIM pp.529-534]. The required bandwidth, however, has been halved.[8]
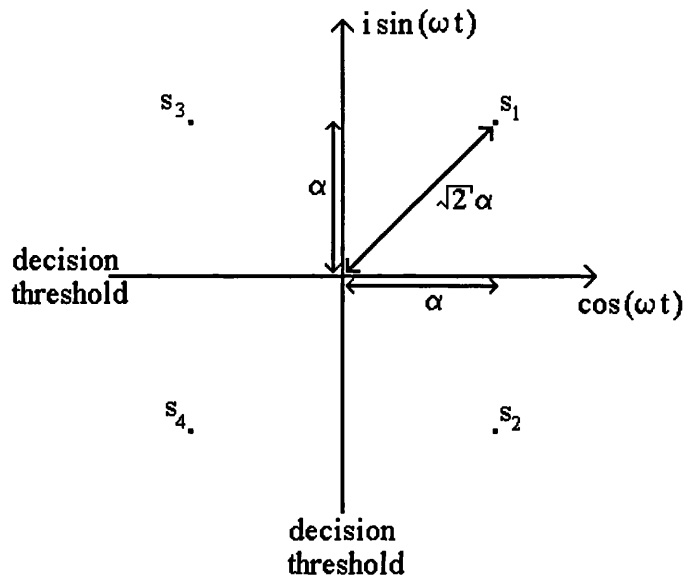
Figure 6.4 -- The signal space representation of a 4-ary phase shift keyed signal constellation with the axes as decision thresholds. Signal s_1 represents '11' and is most likely to be misinterpreted as signal s_2 ('10') or s_3 ('01'), resulting in a single bit error; it is least likely to be misapprehended as signal s_4 ('00')

It is possible to increase the number of number of bits combined in a single symbol and phase shift keyed upon the sin(ωt) and cos(ωt) axes, but this begins to degrade the probability of error performance. This degradation results from the addition of symbols that are *not* orthonormal with respect to each other, figure 6.5a. As the number of bits combined into an M-ary phase shifted constellation increases it is necessary to transmit with greater power to maintain the probability of error performance [TAUB pp.267-269, 472-473][HAMB pp.397-401]. It is possible to

---

[8] in some practical applications the phase discontinuity when the symbol changes can cause difficulties. This can be addressed using the related Minimum Shift Keying (MSK) technique, where the data is smoothed using a sine/cosine wave. A fuller description of MSK and bandwidth comparison with QPSK may be found in [PASU]

ameliorate, but not eliminate, this extra power requirement by simultaneously varying the amplitude and phase: 'quadrature amplitude modulation'. This is amplitude modulation upon a quadrature phase shift keyed constellation, figure 6.5b, resulting in an increased signal separation with a correspondingly lower probability of error but with no increase in the energy required, figures 6.5a and 6.5b [SCHW pp.613-620]. However, M-ary phase/amplitude modulation (with M greater than four) is generally a means to trade energy consumption for reduced bandwidth whilst transferring the same amount of information.
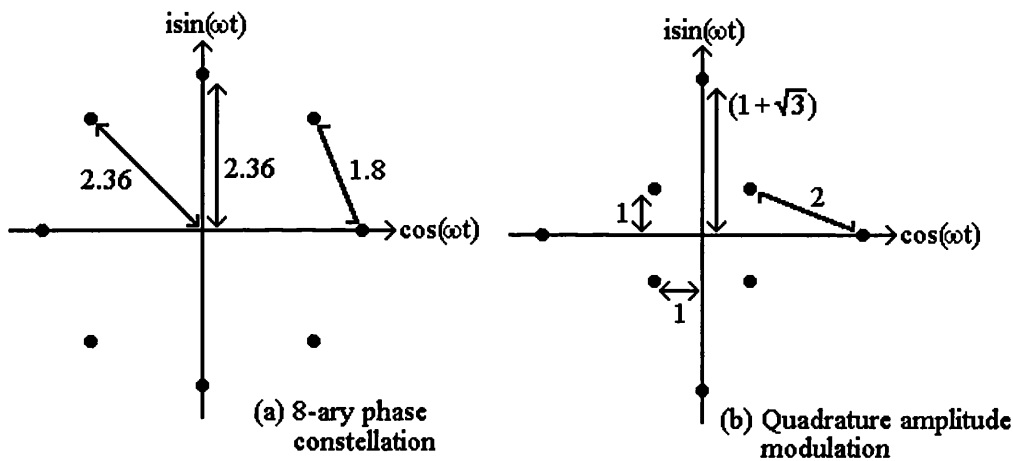
Figure 6.5 -- (a) 8-ary phase shift and (b) quadrature amplitude signal constellations, of equal constellation energy with increased signal separation (shown by the single-headed arrows) of the quadrature amplitude modulated case visible

One can reduce average signal energy whilst maintaining the probability of error performance for a given information set: signal space can be expanded into further dimensions through the use of further orthonormal axes. These can be different frequencies or non-interfering time slots [SCHW pp.620-628]. This increases signal separation and hence improves error performance with no requirement for increased average signal energy [VITEb][STRE pp.651-655]. This reflects the use of a greater bandwidth for transmission.

6.2 Narrowband noise: spread spectrum

Conventional modulation techniques address the probability of error arising from the addition of wideband Gaussian noise to the received signal. They do not, however, provide any protection from narrowband noise interference, such as man-made transmissions. If a narrowband noise source of instantaneous magnitude β is added to a binary signal set of magnitudes $\pm\alpha$ this causes the mean of the received Gaussian probability density function for a low to be $(-\alpha + \beta)$, and the mean value of a high to be $(\alpha + \beta)$, figure 6.6. Without prior knowledge of this narrowband noise the decision threshold, μ_T, would remain that midway between the expected Gaussian means of $\pm\alpha$ (figure 6.1).[9] The increase in the probability of a low being interpreted as a high is greater than the decrease in the probability of a high being interpreted as a low: the overall probability of error has increased.
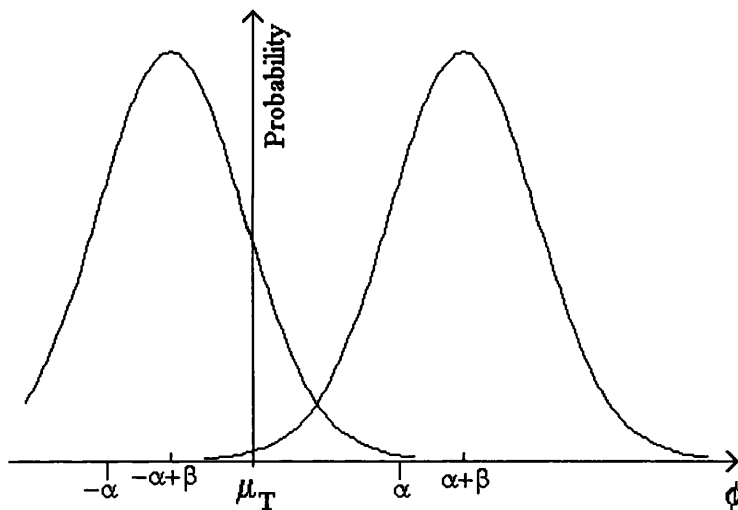
Figure 6.6 -- The probability density functions received when two signals are transmitted with a modulation magnitude of $\pm\alpha$ in the presence of wideband Gaussian noise, with the addition of narrowband noise of amplitude β. Without prior knowledge of this noise the decision threshold, μ_T, does not provide optimum error minimisation

---

[9] in general, the narrowband noise magnitude will not be constant over time and hence changing the decision threshold will not be able to entirely compensate

58

The feedback system is not, in the first instance, expected to be subject to significant narrowband noise (for example, it is unlikely that third parties will wish use a jamming signal to interfere with lecture communication). However, it is *possible* that it may be used in future situations where rejection of this type of noise would be advantageous (such as, for example, the system's use near a commercial radio transmitter). As explored in chapter 5, spread spectrum is a technique that permits many users to simultaneously share the same frequency band, without the problem of catastrophic data loss from collisions between transmissions. The most attractive form of this technique, Direct Sequence CDMA (section 5.4.4), also has appealing narrowband noise rejection, as is proved directly.

A transmitter's data s(t) (with values ±1, for convenience) is mixed with a pseudorandom code c(t) (also ±1) and modulates a carrier at frequency ω, figure 6.7. This is then transmitted through a channel which adds a narrowband interfering signal, also with frequency ω.

Pseudorandom
code
c(t)

Carrier
$\sqrt{2P_s}\cos(\omega t)$

Data
s(t)

Modulo-2
addition

Transmission

Transmitter

narrowband
noise j(t)

Integrator

s(t)

$v_2(t)$

$v_1(t)$

Receiver

$\sqrt{2}\cos(\omega t)$

c(t)
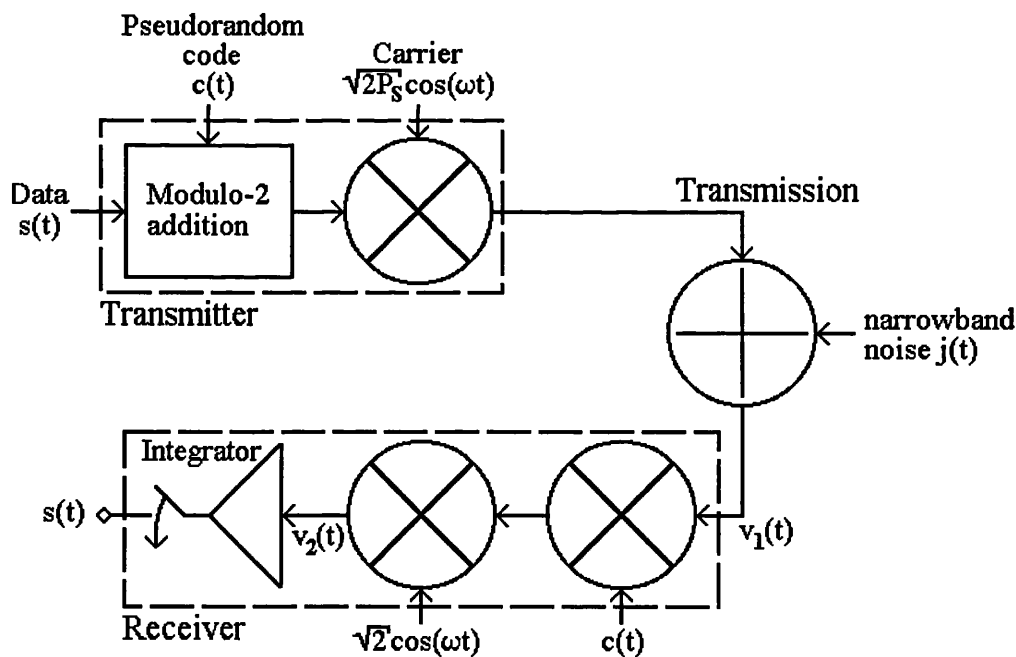
Figure 6.7 -- The generation and reception of a bi-phase shift keyed, direct sequence spread spectrum signal

However, for simplicity, this channel is otherwise perfect with no power fading and no thermal or other noise sources. The receiver's input, $v_1(t)$, is therefore:

$$v_1(t) = \sqrt{2P_S}\ s(t)c(t)\cos(\omega t) + \sqrt{2P_J}\cos(\omega t + \theta) \qquad \text{Eq. 6.10}$$

with the data signal having amplitude √(2P$_S$), and the interfering signal amplitude √(2P$_J$) with some phase θ. An identical, perfectly synchronised pseudorandom code is then mixed with $v_1(t)$ and the carrier re-injected. The signal that is input to the integrator, $v_2(t)$ is

$$v_2(t) = \sqrt{P_S}\ s(t)\{1 + \cos(2\omega t)\} + \sqrt{P_J}\ c(t)\{1 + \cos(2\omega t)\}\cos\theta\}$$

$$- \sqrt{P_J}\ c(t)\sin(2\omega t))\sin\theta \qquad \text{Eq. 6.11}$$

taking advantage of $(c(t))^2$ being unity and applying trigonometric identities. The output of the integration of this signal is maximised if it begins at the beginning of a data bit and has a duration equal to that bit. It is also advantageous if this data bit duration is equal to some multiple of half of a carrier period, as this causes any terms including a $\cos(2\omega t)$ or $\sin(2\omega t)$ component to be zero after integration. The output of the receiver is then

$$v_2(t) = \sqrt{P_S}\ s(t) + \sqrt{P_J}\ c(t)\cos\theta \qquad \text{Eq. 6.12}$$

The information signal has had its bandwidth compressed by c(t); whilst the narrowband noise has had its spectrum widened by the same factor. This widening has an attendant reduction in the spectral power density and so (with bandpass filtration) the effect of narrowband interference on the probability of error is diminished. In general, a DS-CDMA system's reduction of the effect of interference is governed by its 'processing gain':

$$G_P = f_C\ /\ f_D \qquad \text{Eq. 6.13}$$

where f_C is the rate at which the pseudorandom code is clocked, and f_D is the data rate [TAUB pp.724-726]. The direct sequence signal compression averages the

narrowband noise power over the entire region into which the data was spread. This, of course, means that this technique will not improve a system's performance against *wideband* background noise.

6.3 Coding

Returning to the general case, it is possible to reduce the probability of error for a transmitted signal due to wideband noise by appending code bits. Errors in the data may be detected or, by using codes of greater complexity, corrected. Codes can be thought of as providing this error protection in the same manner as previous modulation techniques: adding extra dimensions to signal space. Each transmitted bit can be defined as an orthogonal axis - its transmission occurs during a time period that it does not overlap with, and hence is independent of, every other bit.[10] It should be emphasised that bits are orthogonal *irrespective* of the modulation technique employed in their generation. In the case of a two bit amplitude modulated message, each bit having duration T_B and the message beginning at time zero, these axes are

$$\phi_1 = \sqrt{2/T_B} \ \{A_1\cos(\omega t)\} \qquad \text{where } 0 < t < T_B$$

$$\phi_2 = \sqrt{2/T_B} \ \{A_2\cos(\omega t)\} \qquad \text{where } T_B < t < 2T_B \qquad \text{Eqs. 6.14}$$

The normalisation is provided by the square rooted factor; the term in the curly brackets can be generalised to any modulation type. The provision of further bits in a message *that relate to the original signals* can therefore provide an increase in signal separation in signal space.

In general, a coded data word, n bits long, consists of k data bits and $(n - k)$ appended code bits, and is termed a (n, k) word. These code bits can be included elsewhere in the word; however, as the additive white Gaussian noise in the channel is random and

---

[10] with the further provision that the transmission channel noise at time t_1 is independent of the channel noise at some later time t_2

memoryless (the probability of an error in the ith bit is not affected by any previous bits) the order of the data word and code bits is irrelevant. There are two fundamental approaches to its reception. The receiver can sequentially decide whether each bit is above or below a decision threshold and therefore determine whether each is individually high or low - 'hard-decision decoding'. Alternatively, the receiver can store each received bit and then compare the ensemble result to the expected possibilities. Each bit provides a component vector, with the resultant compared to the theoretically possible vectors in n-dimensional signal space, with the receiver opting for the theoretical vector that is closest to the received one - 'soft-decision decoding'. As might be expected, soft-decision decoding provides superior performance as the calculations are conducted in a signal space with a greater number of orthogonal axes - but requires greater complexity in the receiver [HAMB pp.454-461]. Where necessary, this section assumes the use of optimum soft decision decoding.

In an energy limited system the addition of code bits causes the energy per transmitted bit, data or code, to fall. This increases the probability of a bit error, figure 6.1. Therefore, to be useful, the decrease in the probability of error resulting from the deployment of the code must not only compensate for this increase but exceed it. There is a considerable literature outlining the generation and use of codes extending to complex forms [SCHW pp.636-708]. Only most fundamental codes are within the scope of this work as the information rate from each handset is low (more involved coding techniques tend to be employed for lengthy bit streams), and the complexity of any code-generating circuitry must be limited due to energy constraints.

6.3.1 Repetition codes

The simplest form of code is where the transmitter simply repeats each bit w times to form a (w,1) code. This can be constructed by repeating each bit in turn, or by retransmitting the entire message. Errors may be detected if w is set at two, or corrected if w is three (increasing w further provides greater error protection).

However, repetition coding is not efficient - the energy per bit decrease is substantial, and the rate of increase in error protection less so. It can be shown that for fixed energy transmission it *never* provides a net decrease in the probability of error in an AWGN channel (although it does have utility with channels with other forms of wideband noise, such as Rayleigh).[11]

6.3.2 Combinatorial codes

Greater progress can be made by using the code bits to express the relationship *between* data bits. The commonest means to generate these code bits is through the use of modulo-2 addition of the data bits, which has the truth table:

Input A	Input B	Modulo-2 output
0	0	0
0	1	1
1	0	1
1	1	0

Table 6.1 -- Modulo-2 addition

This modulo-2 addition is applied between the different data bits. As an example, consider a code operating on a three bit data message, that modulo-2 adds bits 1 to 2 and 2 to 3, appending the results to form the data word:

---

[11] Rayleigh noise probability density functions are described in [HAMB pp.371-375]; proof that repetition coding provides no decrease in the probability of error in an AWGN channel, but can in a Rayleigh channel is provided in [ZEIM pp.696-697, 708-712, figure 10.19]

Original data bits: **1 0 0**

Modulo-2 addition of bits 1 and 2: **1**

Modulo-2 addition of bits 2 and 3: **0**

Output word: **1 0 0 1 0**

6.3.3 Error detection

The sequential modulo-2 addition of all of the data bits generates a 'parity bit', creating a (k+1, k) code word (this is an even parity bit; an odd parity bit is the inverse of this). For example:

Original data bits: **1 0 0**

Modulo-2 addition of bits 1 and 2: **1**

Modulo-2 addition of the above result to bit 3: **1**

Output data word: **1 0 0 1**

The receiver repeats the modulo-2 addition on the received data bits, allowing it to determine whether an odd number of errors have occurred. In the above example, if the received data word was 1101 the receiver would know that the data word was erroneous (it is not possible to detect an even number of errors with a parity bit). It is *not* possible for the receiver to correct the error, as it cannot identify the specific bit error; the receiver has to discard the message in its entirety.

6.3.3.1 Code performance

Parity check bits provide powerful error detection. The probability of an undetected error of in an uncoded 11 bit data word, P_{EU}, where the probability of *bit* error is p_U is:

$$P_{EU} = 1 - (1 - p_U)^{11} \qquad\qquad \text{Eq. 6.15}$$

The probability of an undetected error of a 12 bit code word (consisting of that data word with an appended parity bit), with a new, increased probability of bit error p_{EPB} is:

$$P_{EPB} = \sum_{i=2}^{10} {}^{12}C_i \, (1 - p_{EPB})^{(12 - i)} (p_{EPB})^i \qquad \text{where } i = 2, 4... \qquad \text{Eq. 6.16}$$

The use of these equations, in conjunction with those that provide the probability of error for the various modulations, figure 6.1, shows the advantage of parity bit error detection. For example, if the probability of uncoded bit error is 1×10^{-3} using phase shift keyed modulation, the energy per bit divided by the noise spectral density is approximately 9.8dB and the probability of an undetected bit error in the 11 bit data word, equation 6.15, is 1.1×10^{-2}. With a fixed energy budget and constant noise spectral density, the addition of a parity bit causes the energy per bit to fall to 11/12 of its previous value, a decrease in the energy per bit divided by noise spectral density of some 0.4dB to 9.4dB. The probability of bit error has risen to 1.2×10^{-3}; however, using equation 6.16, the probability of an undetected bit error in the 12 bit word has fallen to 1×10^{-4}.

In practice, however, the data is modulated as fast as possible with a fixed energy per bit to reduce the probability of collision (equation 5.1, p.36). Therefore, the parity bit simply extends the transmission *without* reducing the energy per bit - tolerating the increased energy requirement and probability of collision. The probability of bit error at the maximum transmission range of the large class feedback system (chapter 8) is 5×10^{-4} (section 6.3.5). Without a parity bit the probability of an undetected bit error in the data plus address is some 0.6%, equation 6.15. With a parity bit this probability falls to some 0.002%, equation 6.16.

6.3.4 Error correction

Several modulo-2 additions can be conducted between differing sets of data bits - 'block' coding. With the provision of these code bits the receiver can correct errors. For example, the application of the code:

Code bit number	Modulo-2 addition of:
k_1	$d_1 \quad + d_3$
k_2	$d_1 + d_2 + d_3$
k_3	$d_2 + d_3$
k_4	d_3

Table 6.2 -- The generation of a (7,3) code word which can correct single bit errors

to a three bit data word results in the (7, 3) code words

Data bits	Code bits			
	k_1	k_2	k_3	k_4
0 0 0	0	0	0	0
0 0 1	1	1	1	1
0 1 0	0	1	1	0
0 1 1	1	0	0	1
1 0 0	1	1	0	0
1 0 1	0	0	1	1
1 1 0	1	0	1	0
1 1 1	0	1	0	1

Table 6.3 -- The possible (7,3) code words from table 6.2 with a Hamming distance of 3 demonstrating that single error correcting capability

As can be seen, the code words differ in at least three positions: if a single error occurs the data is still recoverable. The difference in the number of positions

between any two words is known as the Hamming distance, d, and the number of errors recoverable in a word, t, is

$$t \le \frac{(d - 1)}{2} \qquad\qquad \text{Eq. 6.17}$$

Block codes do reduce the overall probability of error at a faster rate than the increase due to the reduced energy per bit [SCHW pp.647-648]. They can be generated systematically, with known error correcting properties; this class of codes are known as cyclic codes [MICH]. Further details of these, and the multifarious other code techniques [BLAH][MANS], are not presented here as they do not have great utility for the low-data-rate, low power feedback system.

6.3.5 Coding: system choice

The probability of bit error for the transmitter used in the large class feedback system (system 2, chapter 8) was measured. At maximum operational range of the transmitter, 20 metres, the transmission of 2×10^4 bits resulted in 9 bit errors (a probability of bit error of approximately 5×10^{-4}). This indicates that the probability of an undetected bit error in the 11 bit data word to be 0.6% (section 6.3.3.1). At a range of 10 metres the probability of bit error was less than 1×10^{-4}, indicating a probability of an undetected bit error in a data word of around 0.1% (equation 6.15). If a parity bit is added to the transmission the probability of an undetected error at 20m is 0.002%; at 10m it is 7×10^{-5}% (section 6.3.3.1, equation 6.16).

The system has been designed in the knowledge that it does not need to ensure that every transmission is successful. The addition of a parity check bit will reduce the probability of an undetected bit error in the data plus address by orders of magnitude. However, the feedback system has no network control techniques and tolerates a probability of transmission overlap (collision). The addition of a parity bit would increase the duration of the transmission and increase the probability of collision - with the latter significantly greater than the probability of an undetected bit error. It was decided to concentrate initially on minimising the probability of collision by

minimising transmission duration and tolerate the probability of a bit error in the data and address - no parity check bit will be used. By a similar argument no error correcting codes were used. However, if the probability of collision was such that the probability of an undetected bit error was significant in comparison to the probability of collision it would be profitable to introduce a parity bit.

6.4 Conclusions

Frequency shift keyed modulation has a clear advantage over amplitude shift keying - the feedback system will use the former. Whilst phase shift keying has further promise the system is constrained by the commercial availability of transmitters, and carrier synchronisation in the receiver can prove to be an onerous, and more importantly time-consuming, task. For example, if a phase locked loop is used in the receiver the transmitter must provide a carrier for the time needed for this to settle into synchronicity.

Quadrature amplitude modulation is not attractive as the feedback system is an *energy* constrained system - the efficiency of bandwidth utility is not of concern. The possibilities afforded by M-ary orthogonal signalling are theoretically inviting, but the need for multiple frequencies for transmission (or synchronised handset and transmitter clocks if the orthogonality is provided by timing) in its practical realisation diminishes this.

It is not anticipated that the system will require much narrowband noise rejection - as this is principally an advantage in hostile environments, such as military communication in the presence of jamming. However, this property of spread spectrum modulation increases its advantages with respect to conventional network access methods - particularly as it can be used in conjunction with other error control techniques.

Parity check bits have good error detecting properties. However, the feedback system does not use any network control techniques and the addition of a parity bit

lengthens transmission duration and hence raises the collision probability. It was decided that the feedback system would tolerate the probability of an undetected bit error due to noise and not use a parity bit; for similar reasons no error correcting codes will be used.

Chapter 7 Prototype system

"You should never have your best trousers on when you go out to fight for freedom and truth" Henrik Ibsen

A prototype system was built to provide preliminary data on the utility of feedback. These handsets were intended to gather preliminary data on, and a first insight into, students' use of feedback, and to demonstrate any technical difficulties at an early design stage. The prototype system was used successfully, with the results presented and discussed in chapter 9.

To speed construction, the data modulation, demodulation and clock recovery functions were provided by commercially sourced integrated circuits. Each handset has five buttons. No network access techniques were implemented and no error-correcting or detecting codes were used, to both minimise the duration of a handset transmission and the complexity of the required circuitry. This lack of a network access protocol necessarily leads to a probability of overlap between handsets' transmissions, equation 5.1. Each handset's transmission lasts for 59ms. With the acceptable probability of transmission overlap (and hence data loss) set at 10% and a frequency of one button press per handset every 9 seconds, 8 handsets could be accommodated. An overview of the data flow from a single handset to the central receiver, and on to storage and display by a computer for the lecturer, is shown in figure 7.1.

Data

Handset

Buttons

Data latching

Power control circuitry

Modulation

Address

Transmitter

Wireless transmission

Receiver

Receiver

Demodulator

Memory

Computer

Hard disk

Storage memory

Monitor

Display memory

Parallel Port

Interface
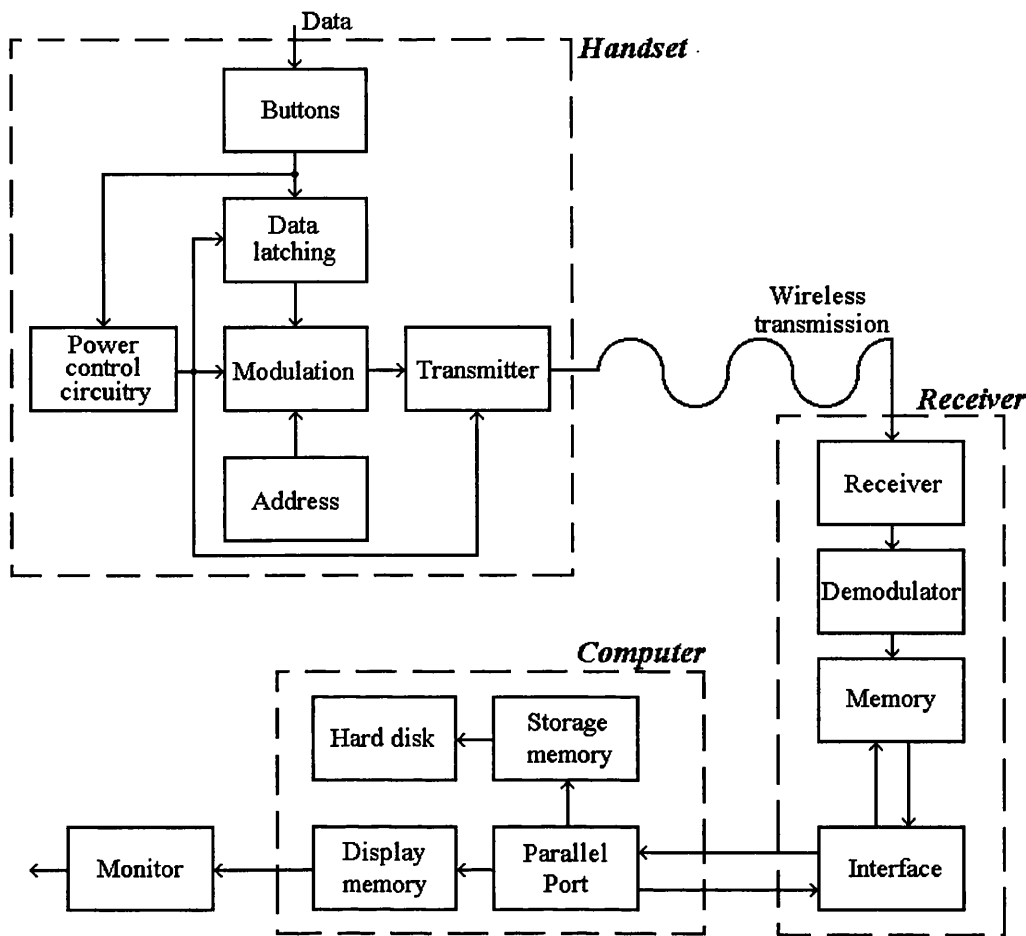
Figure 7.1 -- A schematic of the data flow from a handset to the central receiver, where it is processed by software and stored and displayed by a computer

7.1 Handset

The handset operates on a duty cycle: it is normally unpowered, but when a button is pressed it activates, modulates and transmits the data (appending an address to identify itself), and then deactivates. Each handset transmits immediately after a button was pressed, with a corresponding chance of two handsets transmitting simultaneously with resulting data loss (section 7.4.1). Due to high power consumption (section 7.4.2) the prototype handsets used two 1.2V NiCd rechargeable batteries as a power supply.

The principle of operation of the handset is illustrated in schematic form in the diagrams of figures 7.2a-h, with the active, powered parts of the handset shown in

red. The elements from which the schematic is comprised are further detailed the following sections, with the circuit diagram of the handset included at figure 7.3.

7.1.1 Handset schematic

The 2.4V power supply for the handset is provided by two rechargeable batteries connected in series (marked "Battery" in figure 7.2a and subsequently).
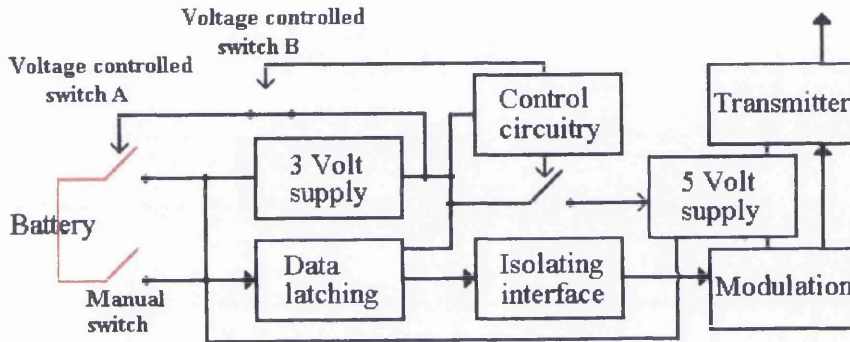
Figure 7.2a -- A schematic of an inactive handset with one of its five buttons marked as 'manual switch'

When one of the five buttons is pressed the battery supply rail is powered, figure 7.2b, supplying the input to a 3V power supply, and applied to the input of a data latching element. The battery supply rail is also connected to an initially inactive 5V power supply.
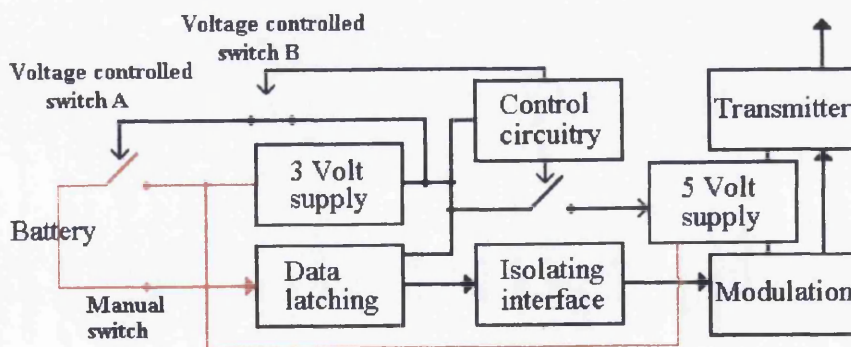
Figure 7.2b -- Immediately after a button is pressed the battery voltage is presented to the input of the data latch and 3V supply

72

The 3V power supply output rapidly stabilises[1] and powers the latch, capturing the data, figure 7.2c. It also provides power to voltage controlled switch A, connected between the battery and the battery supply rail.
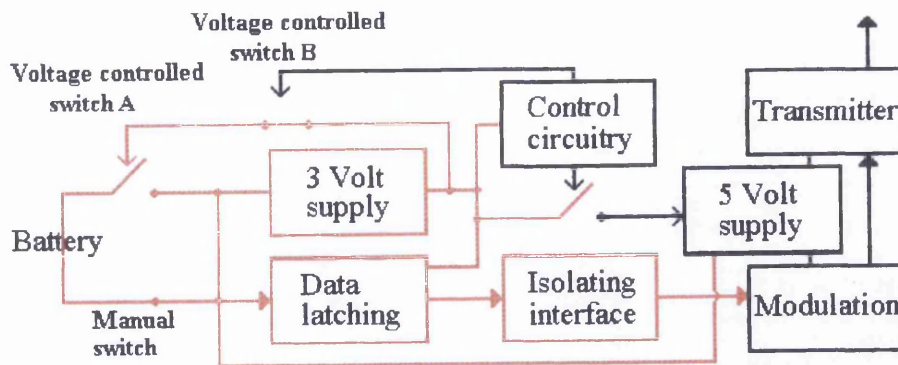
Figure 7.2c -- The handset 3V supply rapidly stabilises and captures the data arising from the button press, and powers voltage controlled switch A

This switch is open when unpowered; when it closes, the operation of the handset becomes independent of the manual switch - there is an alternative route for power to flow from the battery, figure 7.2d.
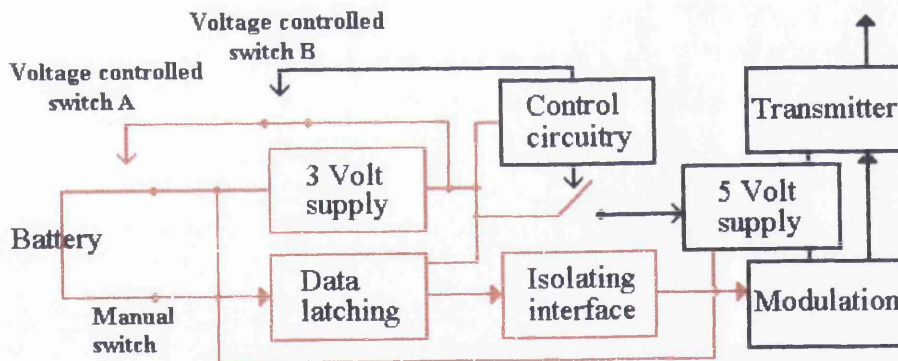
Figure 7.2d -- Voltage controlled switch A closes, providing a battery power supply path independent of the manual switch

The manual switch is later disconnected from the battery supply to the handset, figure 7.2e, leaving the control circuitry as the sole regulator of handset power.

---

[1] 0.2ms after power is provided at its input
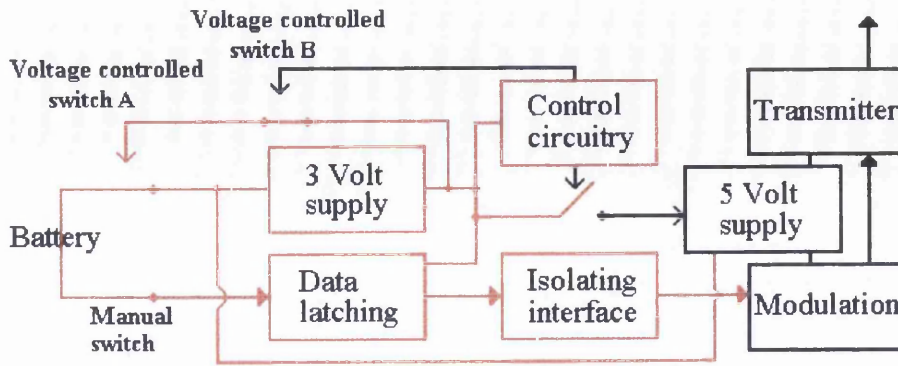
Figure 7.2e -- The manual switch is disconnected from the handset's battery supply

This control circuitry then activates the 5V power supply. This power supply can provide more current, but takes longer to stabilise, than the 3V supply.[2] The 5V supply powers the data modulation, via an isolating interface, so that the data latched at 0V and 3V is modulated with a 'high' of 5V, figure 7.2f. This modulated data is then transmitted.
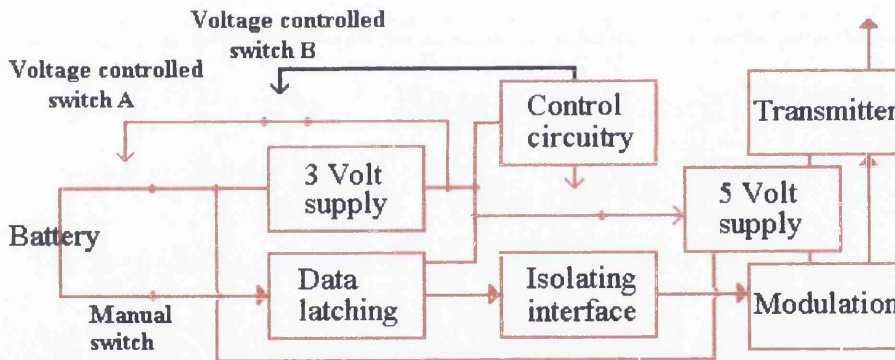
Figure 7.2f -- The 5V power supply is activated by the control circuitry; subsequently the data is modulated and transmitted by devices powered at this voltage

Once the data transmission is completed, the control circuitry begins the process of deactivation, figure 7.2g. Voltage controlled switch B, normally closed, is opened causing voltage controlled switch A to become unpowered and hence also open. Voltage controlled switch B is held open long enough to guarantee handset deactivation.

---

[2] the time to a stable five volt output 1.1ms after power is provided to its input
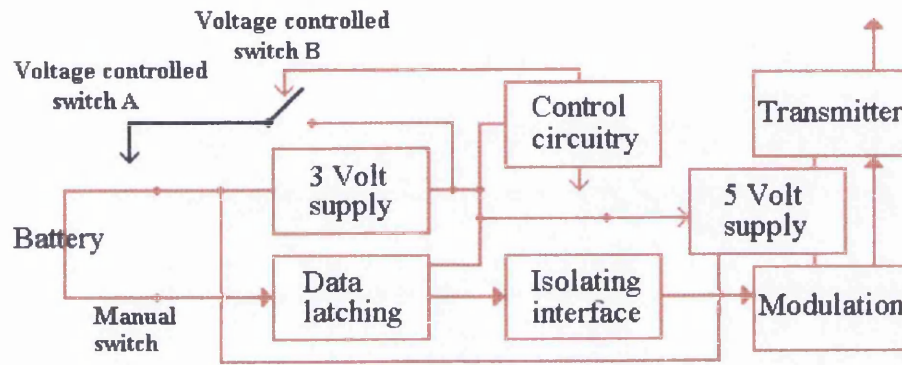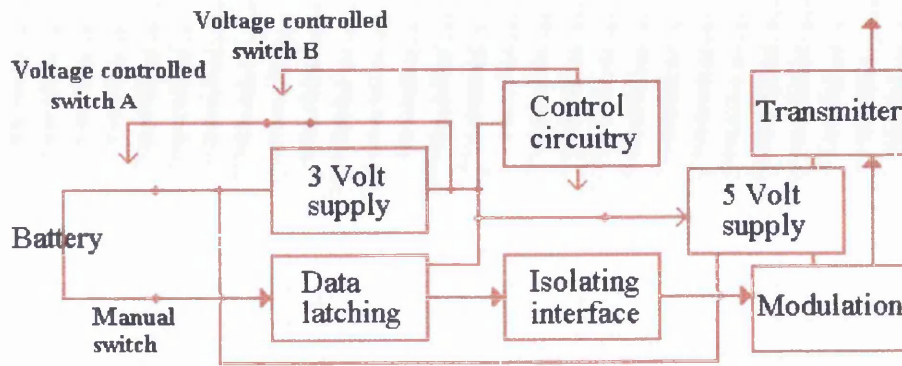
Figure 7.2g -- Handset deactivation via the opening of voltage controlled switch A by voltage controlled switch B, breaking the battery supply circuit

It should be noted that, because of the disconnection of the manual switch from the battery supply rail, a button can remain pressed indefinitely with the handset transmitting the data only once, figure 7.2h. Once the button is released its connection to the battery supply rail is re-established and the handset can be used again.
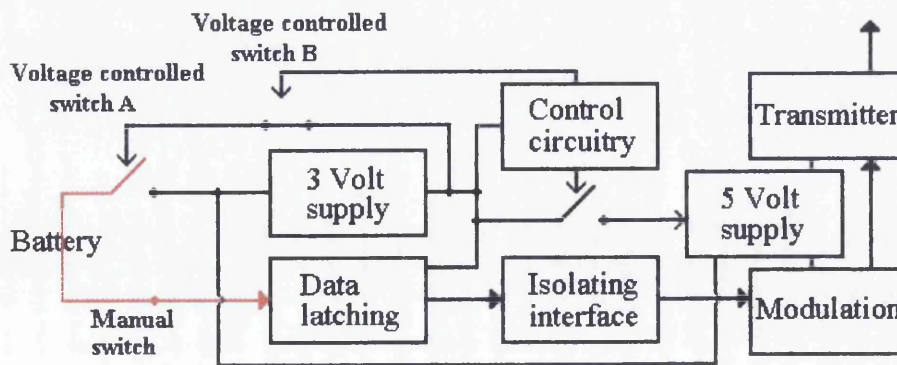
Figure 7.2h -- If a button is held down the handset only transmits once and then deactivates

7.1.2 Handset circuit details

The handset schematic discussed in the previous section is further detailed with comparison to the circuit diagram, figure 7.3.

7.1.2.1 Manual switch

The manual switch of figure 7.2 is one of the five handset buttons (the five switches to the upper left of figure 7.3). The connection to the input rail of the '3V power supply' is provided by transistor T1; this connection is broken by the charging of capacitor C1 through resistors R1 and R2.

7.1.2.2 Data latching

The data latching function is provided by the Set-Reset (SR) latches formed by the usual cross-coupling of the inputs and outputs of NOR gates. Each SR latch is reset via capacitor C7 and resistor R8 to ensure that the latches operate correctly. The active latch's input is connected to the manual switch, with capacitor C2 and resistor R3 providing smoothing. The unused latches' inputs are pulled low by resistor R3.

7.1.2.3 Three Volt supply

A 3V output DC-DC converter[3] which can output up to 5mA of power, stable after some 0.2ms (during which its smoothing capacitance charges).

---

[3] NMP0103 [NMP0]
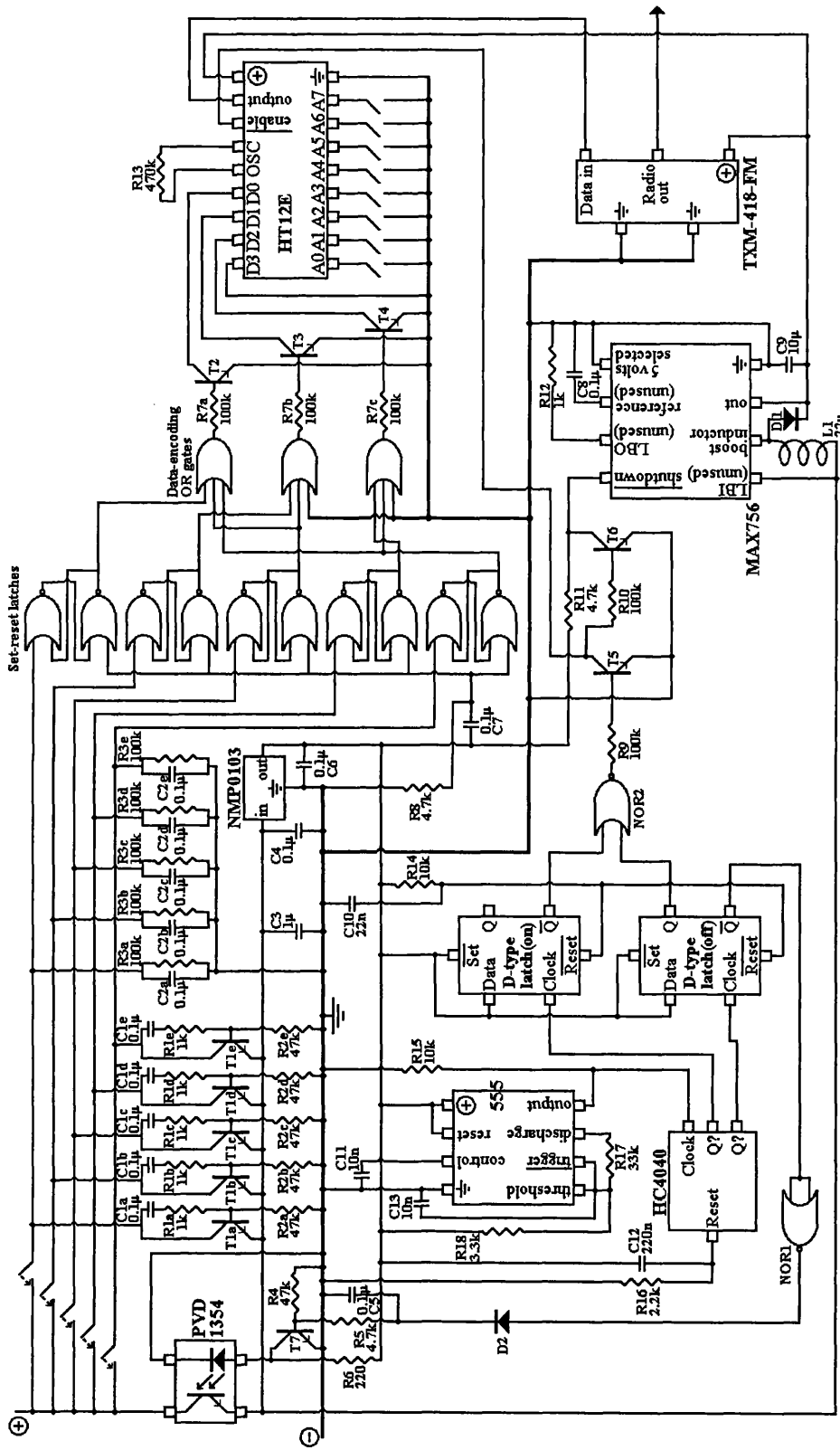
Figure 7.3 -- The handset circuit diagram for the prototype feedback system

77

7.1.2.4 Voltage controlled switch A

A photovoltaic relay[4] provides a route, electrically isolated from the 3V supply, for battery power to both the 5V and 3V supplies. This relay requires approximately 2ms to stabilise in a closed position, depending on the current passed through the light emitting diode (with a diode drop of 1.2V of the 3V supply this current is limited by resistor R6 to 8.2mA). Transistor T7 is used in the deactivation of this switch.

7.1.2.5 Isolating interface

The isolating interface comprises of the encoding of the SR latched data by three triple input OR gates. This data is presented to the modulating IC (section 7.1.2.8) via transistors T2 to T4 (with current limiting provided by resistor R7). This open/closed circuit data format complies with the requirements of the modulator; it also provides a means to isolate the 3 and 5 Volt supplied circuit elements.

7.1.2.6 Control circuitry

The control circuitry is principally an oscillator[5] and a counter. The counter's indicated outputs were latched using standard D-type latches.[6] One latch provides a signal to switch on the 5V supply and to begin the modulation and transmission of the data; the other signal is used to the deactivate the handset.

7.1.2.7 Five Volt supply

DC-DC converter with a 5V output is used to power the transmitter and modulator, capable of providing up to 200mA.[7] This converter requires 1.1ms for stability; it is activated via its 'shutdown' pin (which prevents output voltage gain whilst low).

---

[4] PVD1354 [PVD1]
[5] ICM7555 [ICM7] and 74HC4040 [HC404] respectively
[6] 74HC74 [HC74]
[7] MAX756CPA [MAX7]

7.1.2.8 Modulation

Data modulation and clock recovery were provided using a modulator IC.[8] This CMOS integrated circuit provides each handset with the ability to modulate 8 hardwired address bits and up to 4 data bits. The encoder uses an internal MOS transistor to detect the address/data bits with a resistive connection to the power supply on its input - a high is indicated by relevant pin being open circuit, a low by it being grounded. The prototype uses 3 of the 4 available data bits, and 3 of the 8 address bits.

These bits are modulated using one-third Manchester coding of a master clock, figure 7.4, and are transmitted at the end of a paragogic word, figure 7.5. The maximum oscillation frequency modulator's master clock is 6205 Hz; each paragogic word consists of 73 cycles of master clock, 11.77ms.
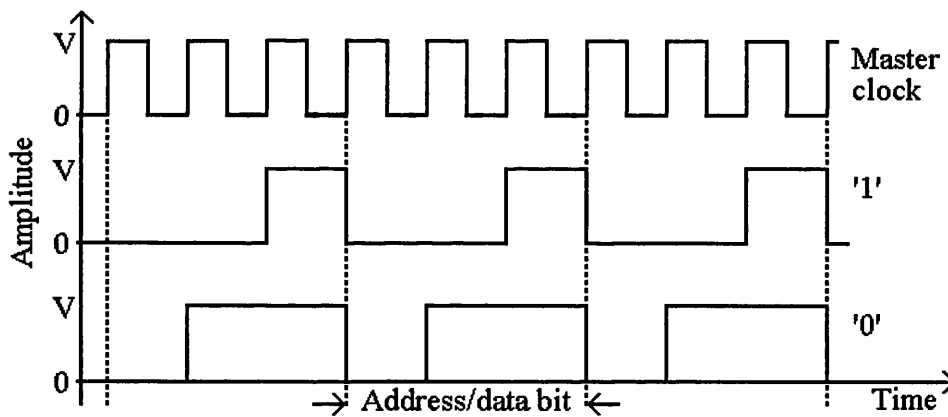
Figure 7.4 -- Modulation uses 3 bits of a master clock per data or address bit
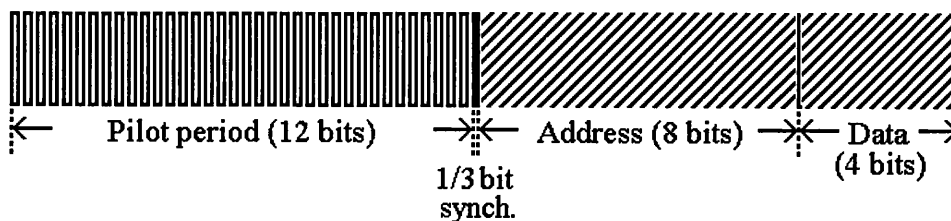
Figure 7.5 -- The modulator's paragogic word with a bit duration of 3 master clock bits: 12 pilot bits, a 1/3 bit synchronisation period and the address/data bits

---

[8] Holtek HT12E [HT12e]

7.1.2.9 Transmission

The output of the modulator was transmitted at 418 MHz.[9] The data is wideband frequency shift keyed [ZEIM, p.178-184] and can have a maximum data rate of 20kHz, resulting in a maximum frequency deviation of 80kHz. However, the modulator limits the prototype handsets' maximum digital pulse rate to 6.2kHz. The transmitter has a minimum line-of-sight operational range of 200m, using a quarter-wavelength monopole antenna.

7.1.2.10 Voltage controlled switch B

The photovoltaic relay requires 2mA of current through its light emitting diode to operate successfully. The schematic 'voltage controlled switch B' deactivates the photovoltaic relay (and hence handset) by reducing diode current to less than this minimum.

After the (fixed duration) data transmission NOR1's output becomes high. This drives transistor T7, connected to the path powering the light emitting diode, into a conductive state. The current passing through this diode then falls below the 2mA minimum. The charge required to hold transistor T7 conductive throughout the process of cutting battery power to the 3V and 5V supplies, and hence NOR1, is held by capacitor C5 (protected from speedy discharge by diode D2).

7.2 Data reception and storage

The radio receiver[10] for the preliminary handset system, figure 7.6, downconverts the data to baseband, with CMOS voltage levels. Eight demodulators,[11] each hardwired with the address of a handset, receive this serial data. When *three* identical copies of the paragogic word from the associated handset are received and matched with the decoder's address its 'valid transmission' pin goes high and the data is presented at

---

[9] by a Radio-Tech TXM-418-FM transmitter [TXM4]
[10] a Radio-Tech SIL-418-FM receiver [SIL4]
[11] a matched Holtek HT12D [HT12d]

its data output pins.[12] The identity of the valid transmission pin is encoded into binary.[13] The valid transmission pin also activates a First In First Out (FIFO) memory device[14] where the data and encoded address are stored. The storage and retrieval functions of the FIFO memory are independent - it acts as a buffer between the incoming data stream and the subsequent download of data to the computer.
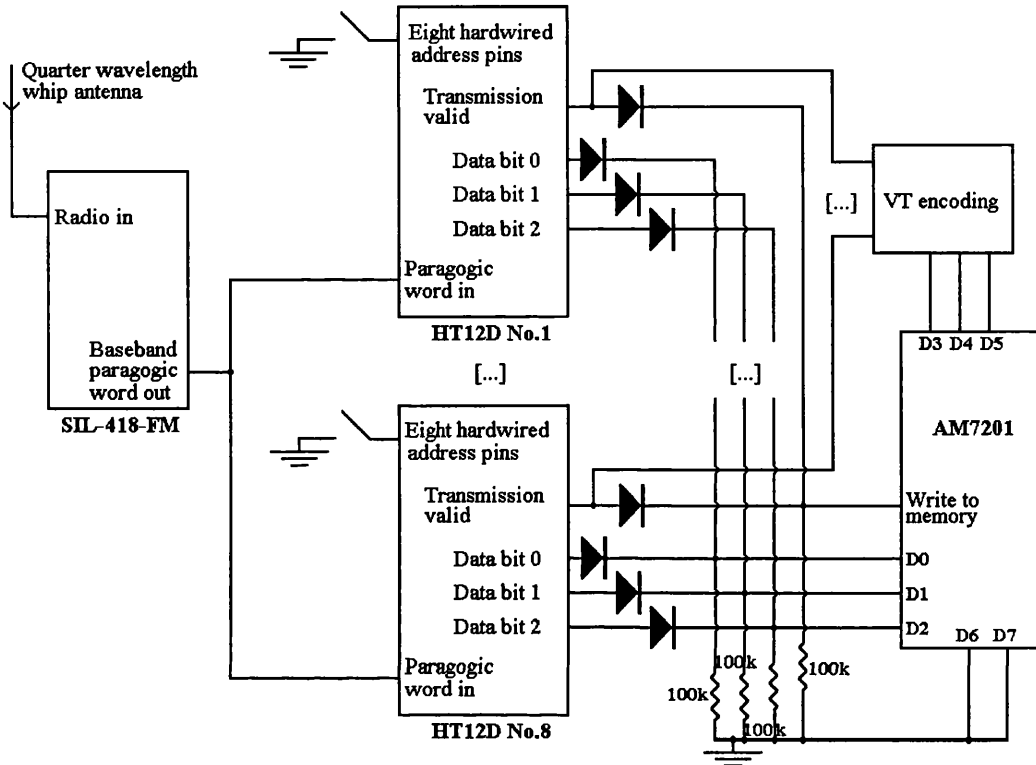
Figure 7.6 -- The data reception and storage circuit

7.3 Data display and long term storage

The data from the feedback system are stored in the FIFO and periodically downloaded to a personal computer (PC). Here they are displayed and stored on the hard disk for later analysis.

---

[12] the requirement for three identical copies to be received is an attempt to provide some rudimentary error protection on the part of the chip manufacturers

[13] using three triple-input OR gates, such that handset 1 is represented by '000' and handset 8 by '111'

[14] AM7201-50RC [AM72]

The data is downloaded from the FIFO storage device into the PC via its parallel (printer) port. A computer program was written to effect this, with a full program listing in appendix 3. The PC uses its status and control pins on its parallel port[15] to interact with the interface hardware, figure 7.7. The program waits until the interrupt pin is sent high by either the FIFO becoming half full or by the clock[16] finishing its timing function (which takes approximately a second). Following this the acknowledge pin is sent high, to ensure that that the 'empty' output of the FIFO gives a true output. If the 'no more data' pin is high indicating there *is* data the program sends 'next byte' high and acknowledge low (if there is no data to download the program restarts the clock and waits for another interrupt signal). The data byte is then read from the data pins, and next byte goes low and acknowledge goes high (ending the FIFO read cycle) and the program again checks the FIFO 'no more data' output. The program continues to read data from the FIFO until it is empty, then restarts the clock and waits for another interrupt signal.
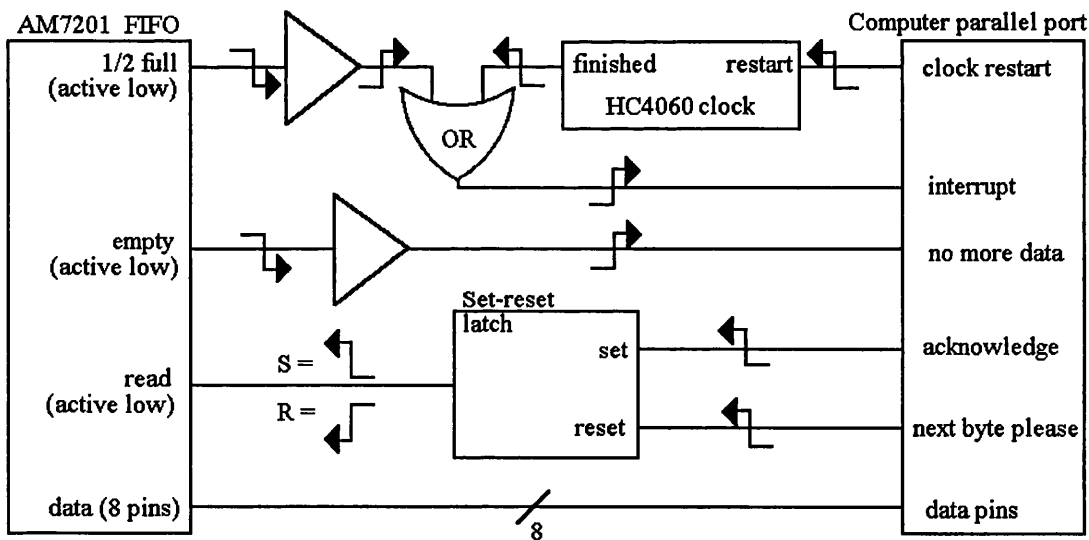
Figure 7.7 -- The circuit diagram of the PC to FIFO interface

---

[15] running in its most basic, bidirectional Standard Parallel Port (SPP), mode of operation. If a later configuration of the parallel port is active, it is changed to SPP for the duration of the programme. This ensures computability for all computers running Windows 3.11 and later.

[16] 74HC4060 [HC406]

As the data is transferred to the computer it is stored to a 'display' memory according to its numerical address (bits D3-D5) - although this memory is designed for an expanded system and can hold up to 256 items - and any pre-existing data in this display memory is discarded. The incoming data (and its address) is also separately transferred to a longer term memory and its time of arrival appended - this memory is permanently recorded on hard disk

7.3.2 Lecturer interface

When the computer program is first activated the FIFO is reset. Subsequently each handset present is required to transmit and hence register its presence. Once all the handsets have been registered (for example, after all the students have arrived) the lecturer then indicates whether a 3-bar (student-prompted feedback) or 5-bar (lecturer-prompted feedback) display is required, figure 7.8.[17] Each histogram bar displayed is a percentage of total number of handsets registered (that is, percentage of the audience). It is possible that a collision could cause a student's registration to fail; the lecturer can easily discern this with only 8 handsets present as the difference in bar height is noticeable. In this case, or at any other time, the lecturer can re-register the number of handsets present. When this occurs, or the display changes between student-prompted and lecturer-prompted feedback, the display is "zeroed" - the display memory, and hence display, is emptied. It is also possible to zero the display via a separate command.[18]

If the three button display is selected, when an opinion is received by the handset it is displayed as the full allocation of the handset. Over the next 30 seconds the height of the displayed opinion decays linearly so that after those 30 seconds the display does not show the handset's transmission. This is to prevent obsolescence in the displayed

---

[17] by pressing "3" on the keyboard for student prompted feedback, or "5" for multiple choice questions
[18] the handset re-registering function is initiated by pressing "r" (in all cases, accidental capitalisation of the letter pressed does not change its function), and the zeroing function by pressing "z". Once the lecture is over the program and FIFO/parallel port interface are deactivated by pressing "e" and hitting the "y" button (this confirmation is used in this case to ensure that the lecturer has not made a potentially time-consuming error)

opinions. It has no effect on the stored data. If the five button display is selected the data display does *not* decay.
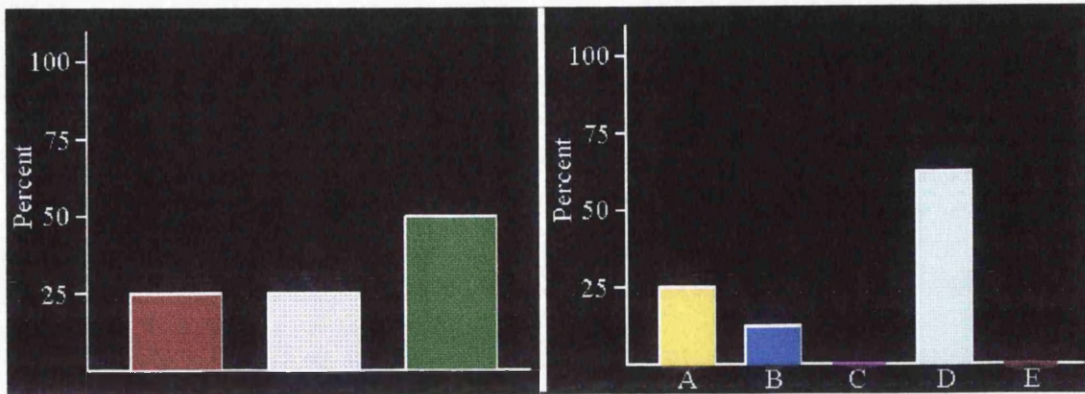
Figure 7.8 -- The 3 and 5 button displays to the lecturer with the total number of handsets present set as 100%. The 3 button display has "speed of delivery too fast" as red, "speed of delivery about right" as grey and "speed of delivery too slow" as green. The multiple choice answers in the 5 button display are labeled as such

7.4 System limitations

This system was designed to be used to provide preliminary experience of lecture feedback - and was successful in that task. However, the number of handsets cannot be expanded to enable feedback to a large-class lecture, and each handset far exceeds the energy consumption compatible with the low maintenance criterion.

7.4.1 Handset provision

Each handset requires 59ms for each transmission (i.e. with the transmission channel utilised), figure 7.9. Six milliseconds are required to ensure the radio transmitter has stabilised, and 7ms subsequent to this to ensure the radio receiver has stabilised reception of this signal.[19] Further, the maximum clock frequency of the modulator is 6.2kHz and this modulator operates by the transmission of *full* paragogic words - each of which requires 11.77ms. Finally, for a successful reception the demodulator

---

[19] the length of this delay is because this receiver is *also* designed to be a low power device - unnecessary in our system

requires three identical paragogic words. These criteria require each transmission to be 59ms long.
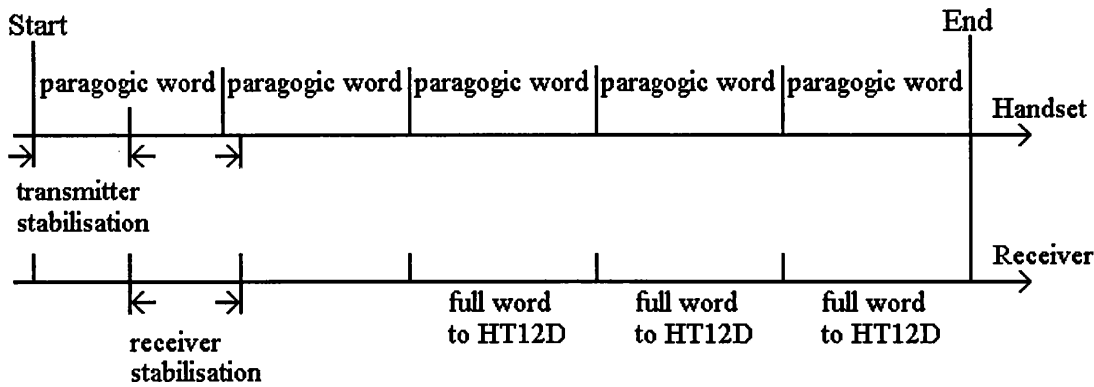
Figure 7.9 -- The time sequence of a handset to receiver transmission indicating the causes of the transmission duration

No network access scheme was used, so transmission duration can be substituted into equation 5.1 (p.36) to determine the probability of overlap with another transmission (collision) for any given transmission. For this preliminary system the acceptable probability of collision was set at 10%, the transmission rate per handset no more than one button press every 10 seconds. With the transmission duration of 59ms the number of handsets that could be accommodated was determined to be 8 - sufficient for a small-class size lecture.

7.4.2 Power drain

On the vast majority of circuit elements constructed from MOS transistors a protection diode is placed between the inputs and the power supply to the device. This provides a discharge path for the accidental application of high potentials to the input (such as electrostatic charges); preventing the destruction of the oxide barrier and hence transistor. These diodes are usually not shown in device data sheets, and can often be disregarded by designers; however, input protection diodes were accidentally instrumental in the operation of the prototype handset.

7.4.2.1 Protection diode: negative effect

An input protection diode was present on the NOR gate inputs used to construct SR latches, figure 7.10, connected to the device's three volt supply. This diode is very briefly in operation between the closing of a manual switch, but before the activation of the NOR gates' power supply, passing the battery voltage directly to the 3V *output* power rail whilst this output rail was at less than the battery voltage. This effect is transient (less than 50µs in duration) and did not appear to interfere significantly in the operation of the handset.

There is a more serious outcome if a button is held depressed continuously. At the end of the transmission of the data the 3V power rail *would not* become unpowered, but present the 2.4V battery voltage to the circuit. As can clearly be seen this would ensure the continued activity of the 'D-type latch (off)', and hence NOR1 would remain powered in a 'high' state and transistor T7 would cause the continuous discharge of current - highlighted in red in figure 7.10. This discharge through resistor R6 represents a *continuous* dissipation of 4mW!

One means by which to block this current discharge path is to introduce a capacitative link in the input path of data to the SR latches. This link enables a 'high' to be SR latched, whilst also preventing current flow, and is marked by the dashed capacitor outline in figure 7.10.

7.4.2.2 Protection diode: white knight

However, this attempted solution revealed that this input protection diode was the reason why the prototype handset was functioning at all.
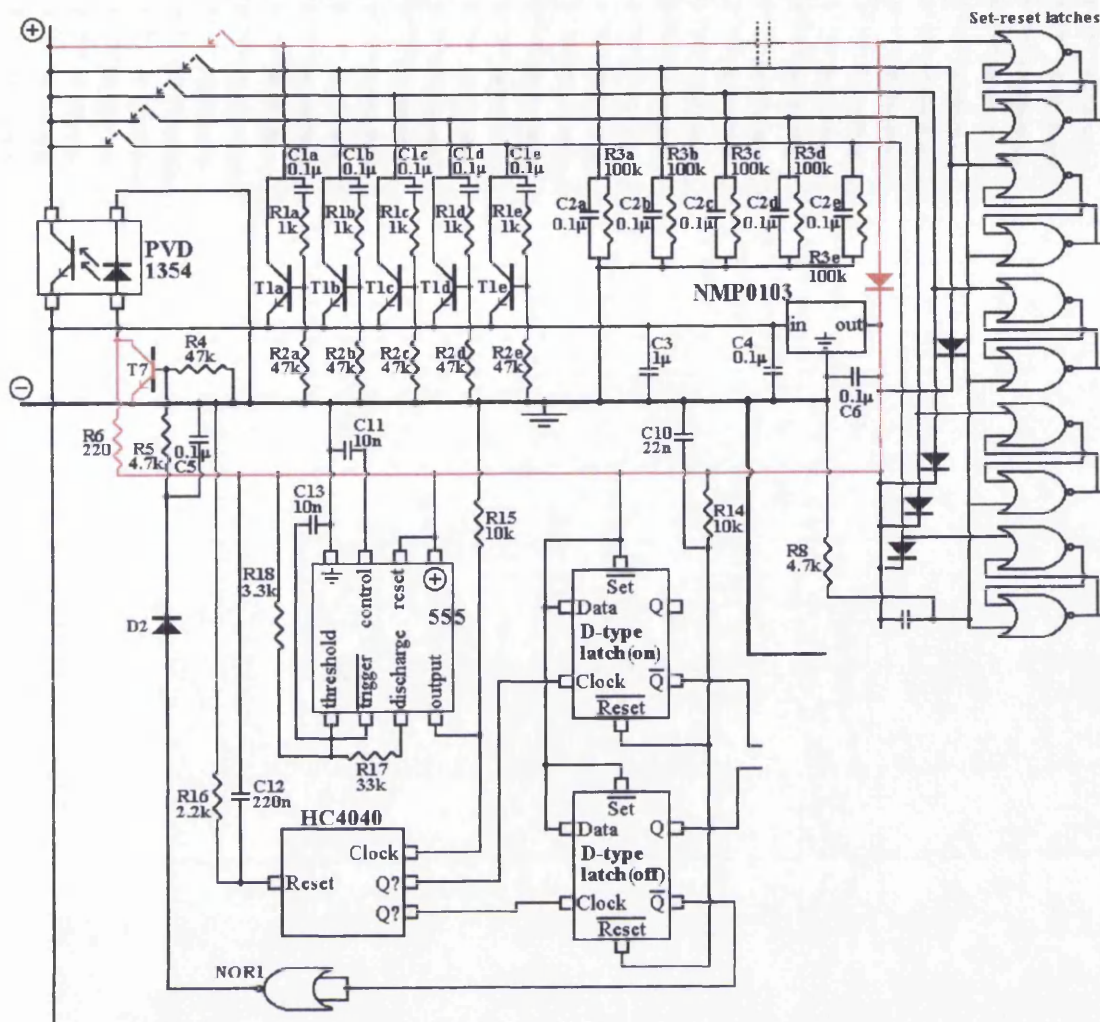
Figure 7.10 -- A section of the handset circuit diagram, figure 7.3, with the power discharge path for a continuous topmost button press shown in red. The capacitor marked by dashed lines was used in an attempt to block this current flow, whilst permitting a voltage signal to pass to the NOR inputs and be latched

The well-known process of DC-DC conversion is the rapid switching of an inductor into and out of circuit. This causes the generation of instantaneous voltages higher than the inductor supply voltage, since the voltage generated is proportional to the rate of change of current flow - high when the switch is operated. This potential is then 'trapped' by a diode; in the case of the handset, diode D1 in the circuit diagram (figure 7.3).

The 5V output DC-DC converter was not intended to activate until directed to do so by the control circuitry. This was to be achieved by a 'switch' (specifically

comprising NOR2, transistors T5 and T6) to hold its shutdown pin low until needed. However, this switch did *not* provide a low to the shutdown pin during the first few tens of microseconds after a handset manual switch was closed. In fact, it was presenting a high, activating, level via the current limiting resistor R11 directly from the 3V supply output rail. The voltage level being applied to the 5V DC-DC converter shutdown pin is high enough to begin activation immediately after the manual button press.

Whilst the diode was providing a direct connection to the battery during the 3V DC-DC converter start-up this was not noticed, as the effect was transient. The 5V supply requires an "inrush" current as it starts operation.[20] This current was provided directly from the batteries and ceased before this connection was broken. When the 5V supply was later started, after the establishment of the 3V supply, the inrush current demand briefly exceeds the maximum supply current for the 3V DC-DC converter (causing a 0.3V drop in the 3V supply) - but the fully charged smoothing capacitance for the 3V supply provides a charge reservoir sufficient to prevent this supply from deactivating. Once the initial inrush current subsides the 3V supply recovers, whilst the 5V supply stabilises.

However, if a capacitor, shown dashed in figure 7.10, was used to block current flow from the battery directly to the 3V supply rail during this period the 3V DC-DC converter has to provide this inrush current *before* it has stabilised. This causes a voltage drop as the 3V smoothing capacitance is not fully charged for 200µs and so has a limited ability to meet the inrush current demand. This voltage drop continues until the voltage on the 5V DC-DC converter shutdown pin is less than that necessary for this converter's activation (approximately 1.35V). The voltage provided by the 3V DC-DC converter then begins to rise. However, as soon as this output exceeds 1.35V the inrush current begins again - and the 3V supply output voltage drops once again. The maximum supply current for the 3V DC-DC converter's output (5mA) only satisfies the initial charge demand (that is, inrush current demand over a greater time) of the 5V supply if the button press duration exceeds 0.4 seconds.

---

[20] during the first 10µs of operation the 5V DC-DC converter has a high current demand as the charge is pumped by the inductor into "storage" behind diode D1 - this current exceeding 20mA at its height

Therefore, the excess handset power consumption if a button was held down for lengthy periods, via the NOR gates' input protection diodes, was tolerated in the prototype feedback system to permit its successful operation. However, with the problem thoroughly characterised it can be solved in subsequent handset development.

7.5 Conclusions

It was felt crucial that the project have some experience with lecture feedback before proceeding to an attempt at providing a large-class feedback system. To this end, a prototype feedback system was constructed. To speed its development construction and development a commercially sourced matched pair of integrated circuits provided the data modulation/demodulation and clock recovery functions. This proved to be non-ideal: the system could not be expanded to a size congruent with that of a large lecture. The theoretically maximum bit rate of the modulation was 6.2kHz; however, the specific technique used by the ICs increased the duration of handset transmission to 59ms. With comparatively lax collision criteria - that no more than 10% of total transmissions overlap and each handset being used once per 9 seconds - 8 handsets could be accommodated. This was acceptable, as with such a low number of handsets it was easily possible to observe if, for example, any handsets had failed to respond to a multiple choice question.

It was difficult to minimise excess handset power consumption. In the event of a prolonged button press a continuous power consumption of 4mW occurred. This difficulty was, however, fully analysed and understood and was amenable to improvement in later developments.

The prototype handset permitted the observation and initial exploration of the utility of lecture feedback (explored more fully in chapter 9). The technical experience in the construction and operation of a feedback system permitted the development of the much more potent handset discussed in chapter 8.

Chapter 8 Large-class feedback: system 2

"Plus, he wanted to go somewhere the Culture had never been, and well, explore"

Iain M. Banks, 'Excession'

The prototype system has two principal inadequacies: the maximum handset complement was only 8, and the handsets dissipate 4mW continuously whilst a button is held down, which could easily result in rapid draining of the batteries. These problems were addressed by introducing a new modulation/demodulation system, and conducting a redesign of the handset power-up procedure.

To tolerate the entire audience transmitting within a short period of time (such as when they are asked a question that can be answered immediately) the maximum rate of student button presses was increased to 1 transmission per handset per second. The transmitter used in the prototype system has a maximum bit rate of 20 kilobits per second. Therefore, the minimum duration of each handset's eleven bit transmission would be 0.55ms.[1] In a system comprising of 256 handsets, each generating 1 transmission per second, the probability of collision was calculated, using equation 5.1 (p.36), to be an unacceptably high 25%. It was concluded that it would be essential to use a transmitter capable of modulation at a much higher bit rate in the system 2 handset.

This chapter details the principle of operation of system 2 handset transmissions, their subsequent reception and storage in a memory device. The download from this memory device to a computer for display and storage is identical to the prototype system. System 2 fulfills the key design criteria - low cost; minimal maintenance; capable of use in very large class lectures. A system 2 handset and receiver were constructed and messages were successfully transmitted.

---

[1] this is an absolute minimum, as it disregards the need for preamble that is required for a receiver to successfully lock on to the transmitter's signal

8.1 Start up, data capture and handset power control: board 1

The principle of operation of the data capture and power control in the system 2 handset is largely similar to that of the prototype system. The start up sequence is, however, modified to prevent the power dissipation which would otherwise be experienced if a button were pressed continuously. This is effected by capacitatively blocking the button press from causing power flow through the input diode of the data latches, and applying this battery supply directly to the output of the 3V supply during the inrush current of the 5V supply so that it can be accommodated.

The block diagram of the handset in its inactive state is shown in figure 8.1a; as before the "manual switch" is one of the five student operated buttons.
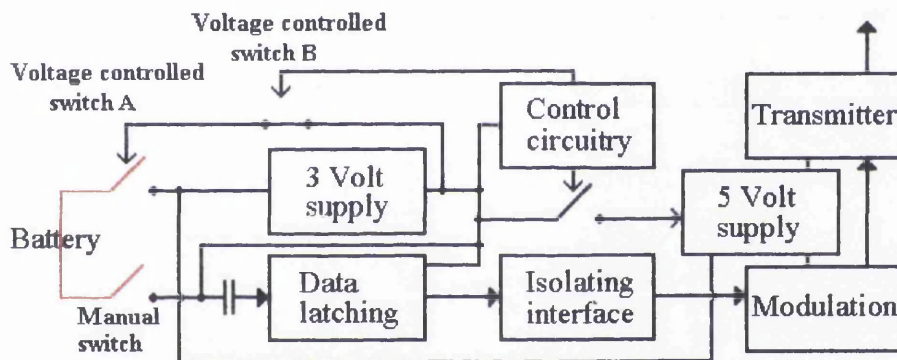
Figure 8.1a -- A block diagram of the handset in its inactive state

The energy dissipated by each system 2 handset transmission requirement is very much less than that of the prototype system, section 8.9, but it is not low enough to enable the use of photovoltaics as a power supply, section 4.3.2. The most appropriate power supply were 2 alkaline 'AA' batteries, which have an energy reservoir of some 15kJ [DURA]. The 3V supply has been retained in this design to permit these batteries to be easily replaced by two 1.2V NiCd rechargeable batteries, acting as a power reservoir for a photovoltaic power supply, once the energy dissipated per transmission has been sufficiently reduced.

When the manual switch is closed the battery voltage is applied directly to the *output* of the 3 Volt supply[2] and a 'high' is presented to the data latches, via a capacitative link; figure 8.1b. The capacitative link ensures that there is no current flow through the input diode of the data latches (as outlined in section 8.4.2) in the event of a prolonged button press. As the 3V supply rail is initially connected directly to the batteries the transient inrush current to the 5V power supply[3] (section 8.4.2.2) can be accommodated.
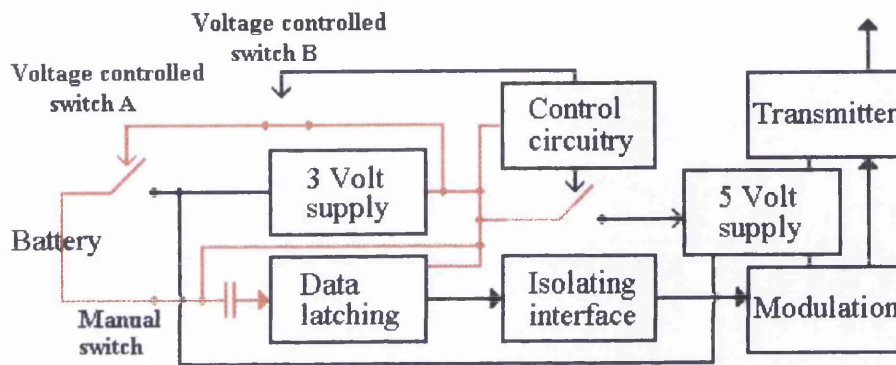
Figure 8.1b -- Power and signal flow resulting from a button press

The battery voltage then closes voltage controlled switch A.[4] This provides a route for battery power separate from the manual switch, figure 8.1c; the data is latched.
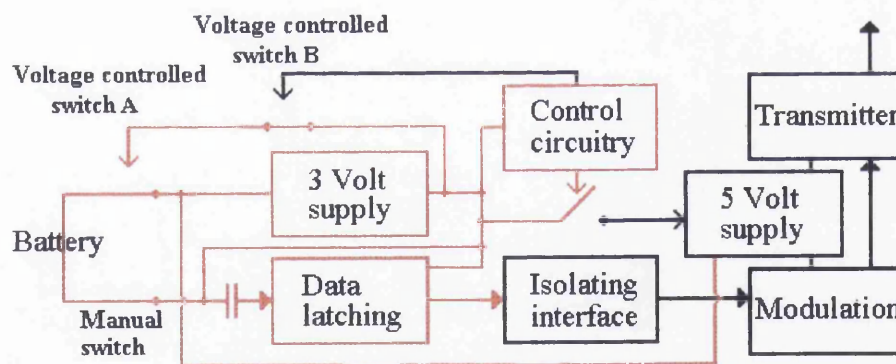
Figure 8.1c -- Voltage controlled switch A closes, establishing a connection from the battery to the 3V supply separate from the manual switch; the data is latched

Subsequently the capacitative link to the data latch charges, causing the voltage level at its input to drop to 'low', and the manual switch can be opened, figure 8.1d.

---

[2] NMP0103S as in the prototype [NMP0]
[3] MAX756CPA as in the prototype [MAX7]
[4] PVD1354 photovoltaic relay, as in the prototype [PVD1]
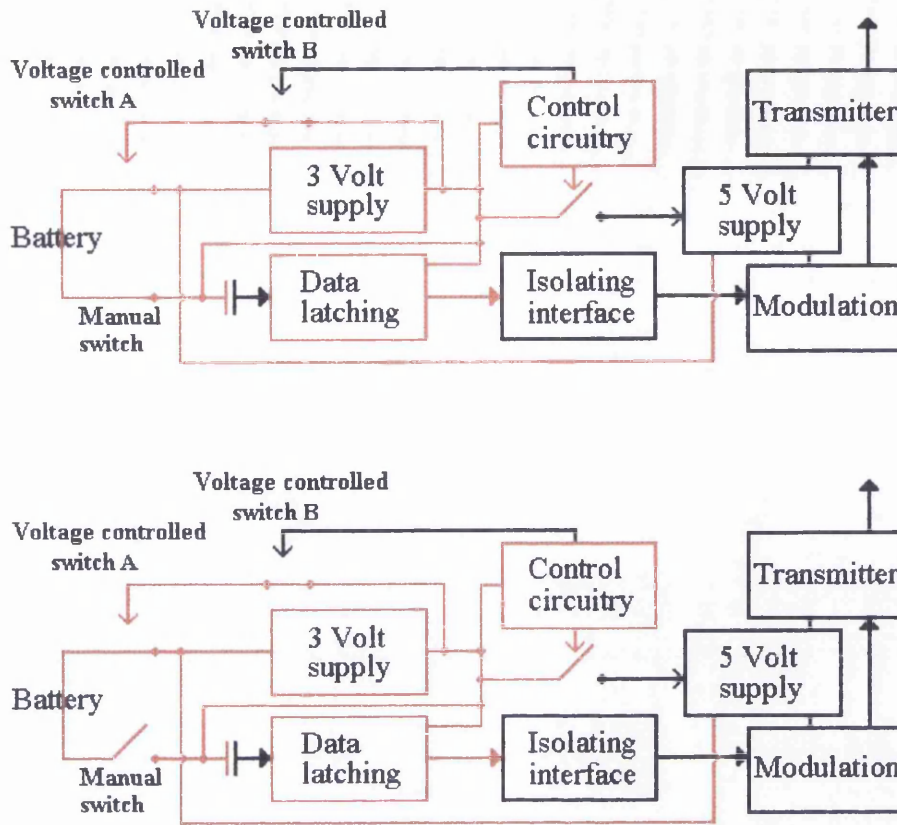
Figure 8.1d -- The capacitative link charges to a level such that the voltage at the input of the data latch becomes low; the manual switch opens leaving the connection via voltage controlled switch A to power for the 3V supply input

The control circuitry then activates the 5V supply; the data is modulated and transmitted with an identifying address (detailed in section 8.2). Once this has been completed, the control circuitry receives an input from the data modulation and transmitter control circuitry, and opens voltage controlled switch B. This switch is held open such that voltage controlled switch A opens and the handset returns to its inactive state, as in the prototype handset. The circuit diagram for the start up, data capture and power control functions - board 1 - is shown at figure 8.2; the parts list for figure 8.2 is included as appendix 11.
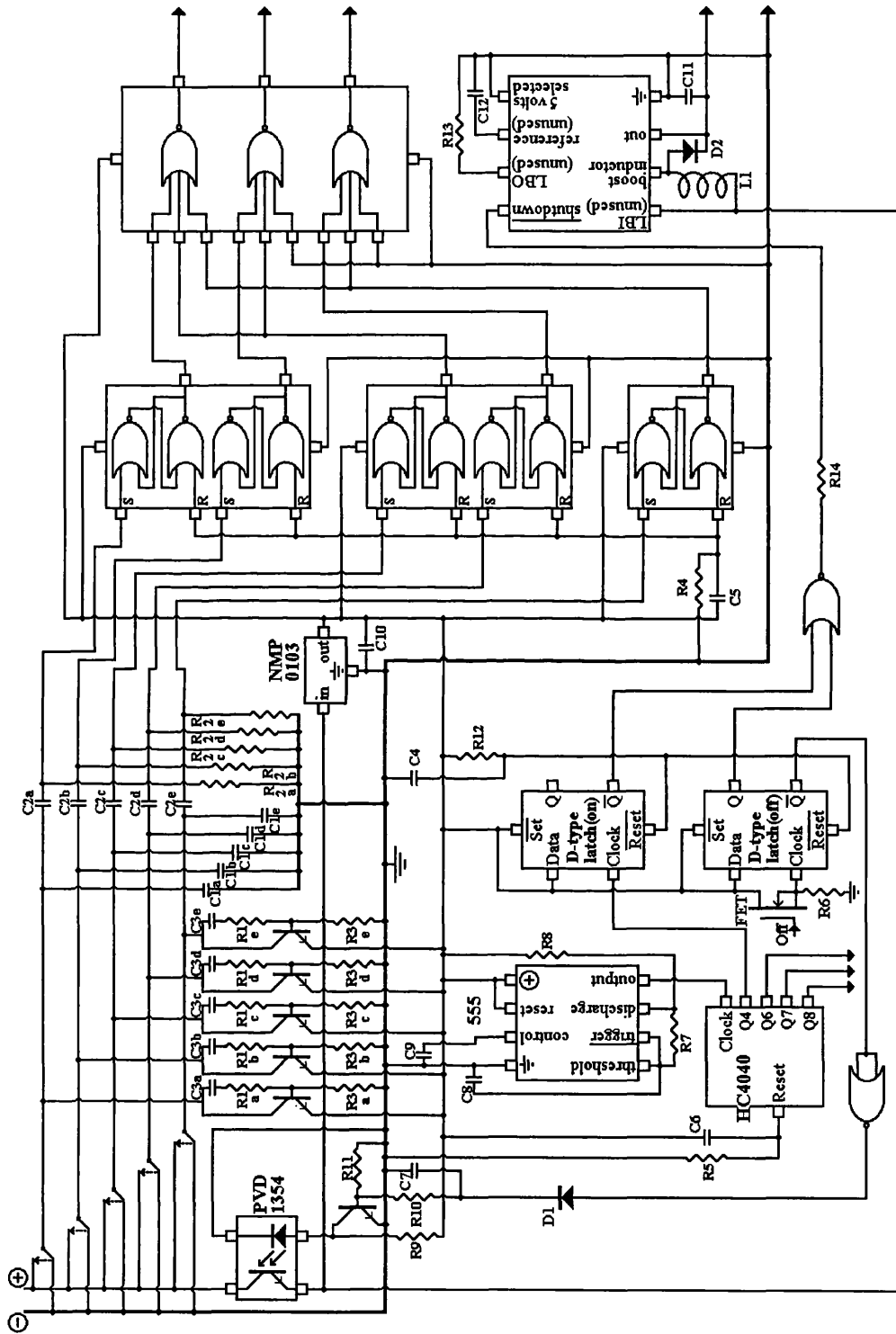
Figure 8.2 -- The circuit diagram for board 1

8.2 Data modulation and transmitter control: board 2

The system 2 handset data modulation function is substantially different from that of the prototype handset; a block schematic is shown in figure 8.3a. When the 5V supply is activated on board 1 the counter, oscillator and parallel-to-serial converting shift registers are powered.[5] Inputs "1" and "2" to this board are from the control circuitry on board 1.[6] Whilst the oscillator is stabilising, taking 9.5ms to do so,[7] the power dissipation in other parts of the system should be minimised. Therefore during this period both the shift register and counter clock inputs are disabled and hence inactive; additionally, the shift registers have not loaded the data from the parallel inputs. The transmitter is unpowered for the first part of this oscillator stabilisation.

Figure 8.3a -- Power-on state of the modulation circuit during the initial oscillator stabilisation

---

[5] a 74HC4040 [HC404]; an HA7210 oscillator module using a 1MHz crystal oscillator [HA72]; and five 'daisy chained' 74HC165 chips [HC16] respectively

[6] input 1 goes high when outputs Q6 and Q7 of the board 1 HC4040 go high; input 2 when outputs Q7 and Q8 go high (where the notation QX indicates a count of 2^X cycles of its 32kHz clock input)

[7] its 5V supply does not complete its power-on stabilisation for a further millisecond after its activation, but the oscillator module stabilisation delay is insensitive to the supply variations experienced

After 5.6ms input 1 goes high, closing the switch to provide power to the radio frequency transmitter and output amplifier, figure 8.3b.[8] The FM transmitter requires 3.9ms before its output is stabilised at 916.36MHz; during this period the modulation input is held open circuit[9] and the output amplifier is off to minimise power radiation from the aerial (further detail on the radio transmitter is included at section 8.3.1).
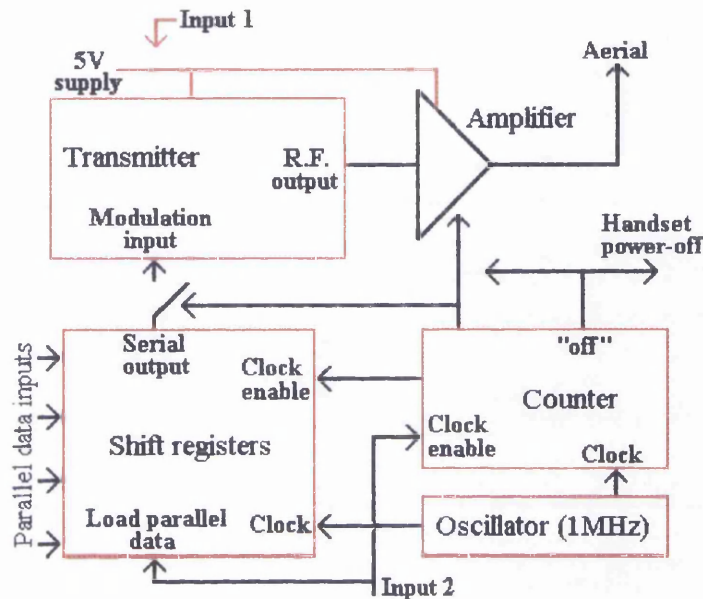
Figure 8.3b -- After 5.6ms input 1 goes 'high', providing power to the transmitter and power amplifier. The transmitter requires 3.9ms to stabilise; during this period the radiated power is minimal as the power control input of the amplifier is held 'low' by the counter

The transmitter and oscillator are both stabilised 9.5ms after the 5V supply activation; their stable outputs are represented in purple in figure 8.3c.

Input 2 then goes 'high', activating the counter clock input and causing the shift registers to load the data from their parallel inputs, figure 8.3d. This parallel data for the shift registers consists of, and will be sequentially output as: preamble, clock synchronising bits, handset address bits and 3 bits encoding the identity of the button that was pressed. Control of the modulation board and transmitter now passes to the onboard counter.

---

[8] this switch is also PVD1354 photovoltaic relay [PVD1]

[9] by a DG419 analogue switch - open circuit being a resistance of least $10^{10}\Omega$ [DG41]
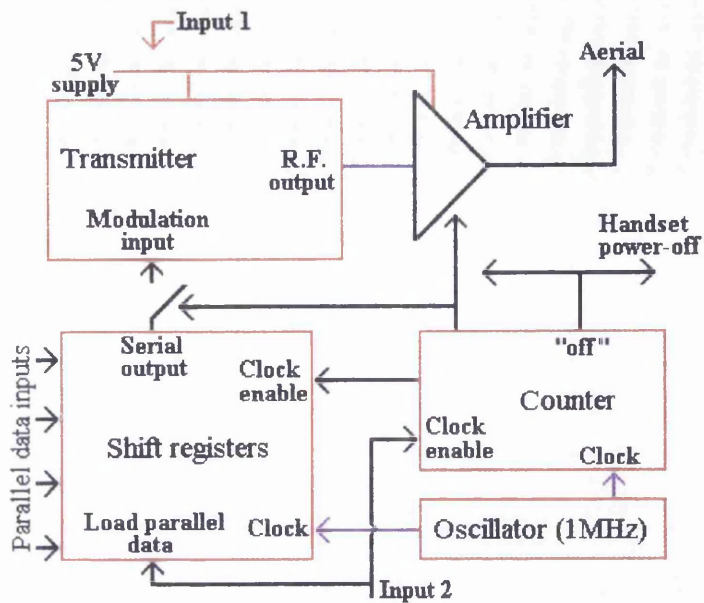
Figure 8.3c -- 9.5 ms after the 5V supply is activated the modulation board oscillator and transmitter have stabilised (shown in purple)
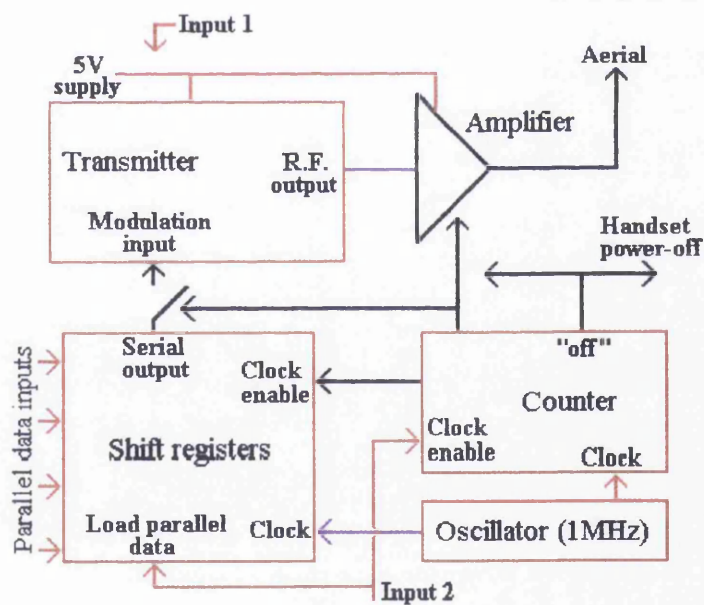
Figure 8.3d -- Input 2 goes 'high', activating the counter's clock input and loading the parallel data into the shift registers; control of board 2 and the transmitter now passes to the onboard counter

The counter increments by one every falling edge of the 1MHz clock. On the twenty-fourth falling edge it outputs a "modulation go" signal, figure 8.3e. This causes: the

amplifier to activate, the switch on the transmitter data input to close and the shift register clock input to activate. The shift register clock serially outputs one bit of the stored (parallel input) data on each *rising* edge of the 1MHz clock. Therefore, the first transmitted bit (a '1') only has a duration of 0.5µs; thereafter, the transmitted data rate is one megabit per second (1Mbps^{-1}). This arrangement is to ensure that the time averaged transmitter frequency deviation oscillates about zero (section 8.3.2)
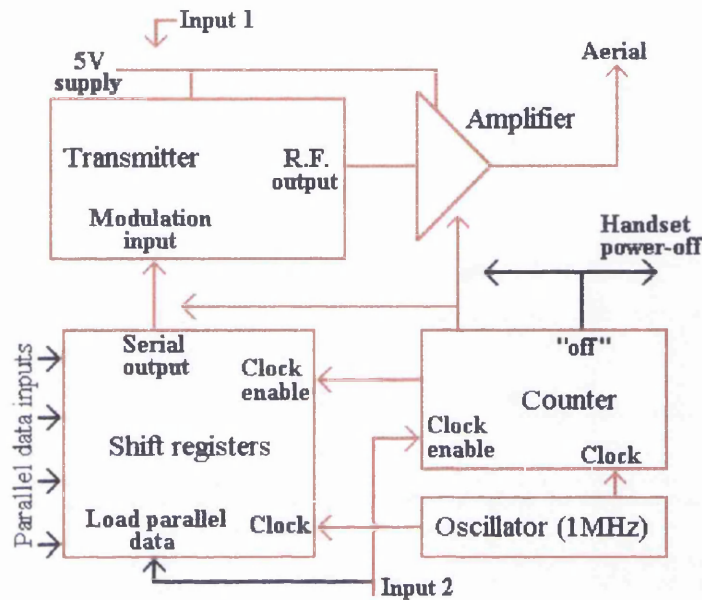
Figure 8.3e -- Transmission begins after the counter reaches 64

On the sixty-fourth falling edge (0.5µs after the end of the last transmitted data bit) the counter outputs an "off" signal, figure 8.3f. This signal is transferred to the "off" input of board 1 and the handset begins to power down. To ensure minimal radiation from the aerial during the handset deactivation process (that is, to minimise the duration of the handset radio transmission) the counter "off" signal also powers down the amplifier. The detailed circuit diagram for board 2 is shown at figure 8.4.
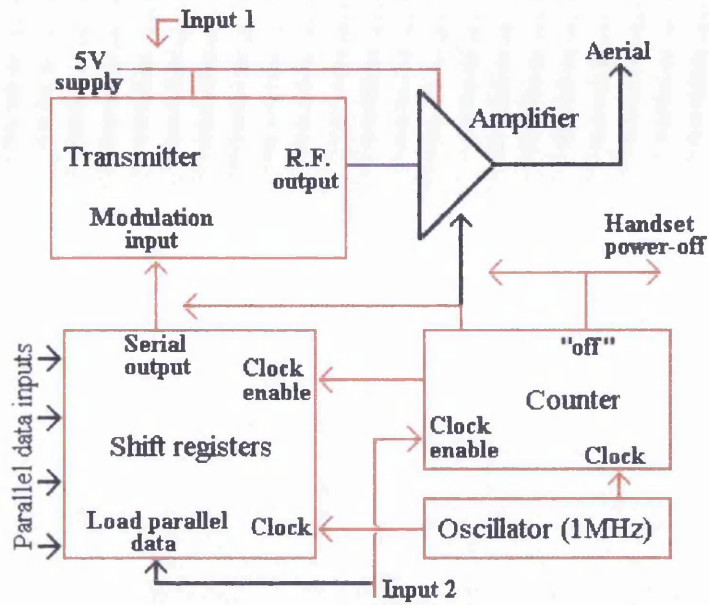
Figure 8.3f -- After transmission an "off" signal is generated by the counter,
deactivating the amplifier and transferred to board 1 to power down the handset
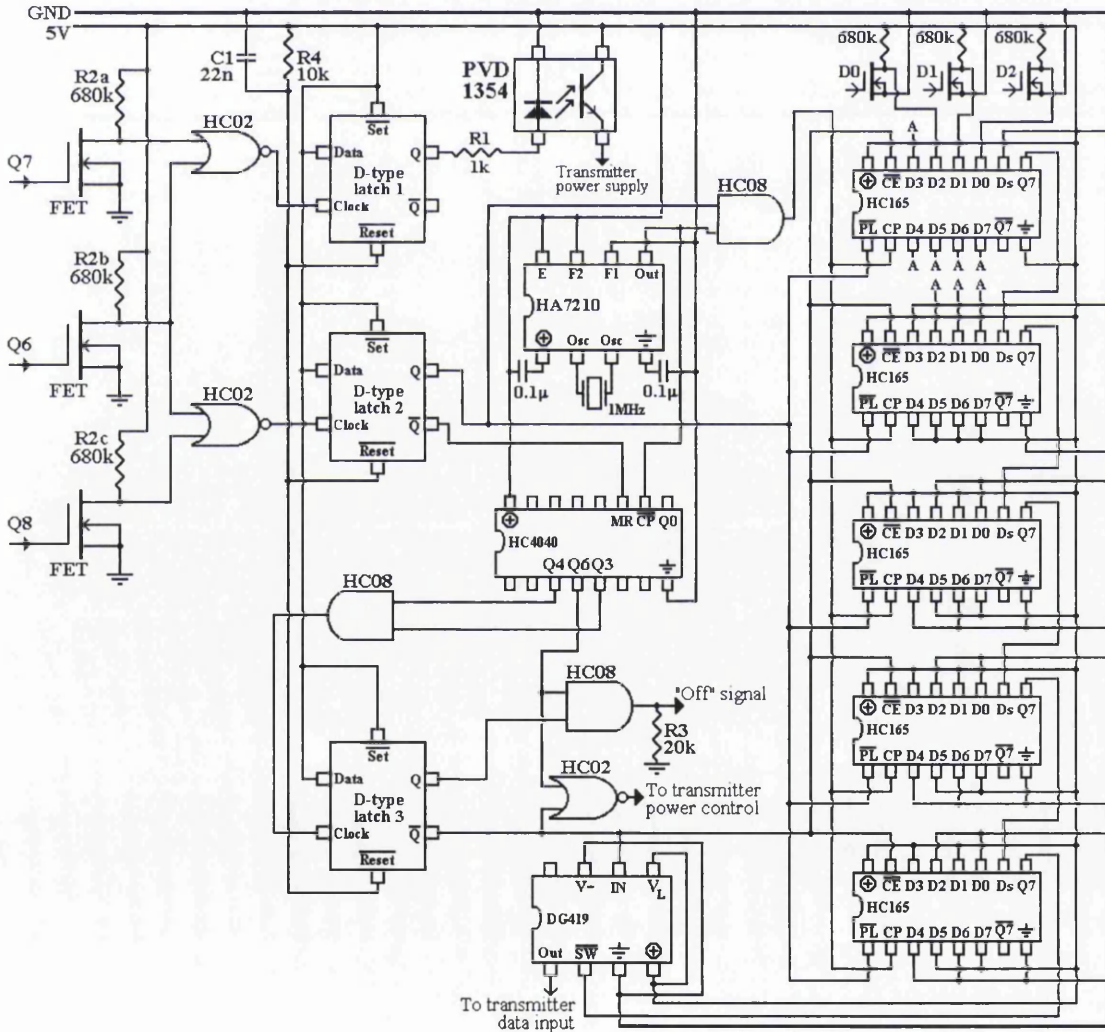
Figure 8.4 -- The circuit diagram for board 2

8.3 Radio frequency transmitter and receiver

A commercial radio frequency transmitter[10] is used to provide the upconversion of the data from baseband to 916.36MHz. Its associated receiver[11] downconverts the data back from radio frequency to baseband; its output is at CMOS voltage levels. This section briefly outlines the principle of operation of these chips (circuit diagrams are included in appendix 4). Reliable transmission under expected operational (lecture theatre) conditions is obtained within a range of 20m.

8.3.1 Transmitter

The transmitter uses frequency multiplication to generate a stable carrier from a lower frequency crystal oscillator. A voltage controlled oscillator (VCO) is placed in a phase lock (feedback) loop where its output is compared in phase to a reference frequency, figure 8.5 (where for simplicity the modulation input is assumed to be held open circuit). The frequency output, F_{OUT}, from the VCO is divided by 32 to give F_{COMP}; the reference frequency, F_{IN}, is a 14.31818MHz crystal oscillator. The phase detector outputs a voltage level proportional to the instantaneous phase difference between F_{COMP} and F_{IN}. This voltage is then applied, with its high frequency component removed by a low pass filter, to the VCO. The VCO output frequency is therefore forced, over some milliseconds, to have a frequency 32 times greater than, and phase identical to, that of the reference. It is important to note that the *phase* detector changes the VCO *frequency* - which is the time derivative of phase. Therefore, the VCO acts as an integrator of the phase difference between F_{IN} and F_{COMP}. An increase in the upper limit of those frequencies *not* passed to ground increases the speed with which the VCO will lock to the reference signal, as lower frequency phase differences will have greater aliasing of the phase difference between F_{IN} and F_{COMP}.

---

[10] RF9901 [RF99t]
[11] RF9902 [RF99r]

100
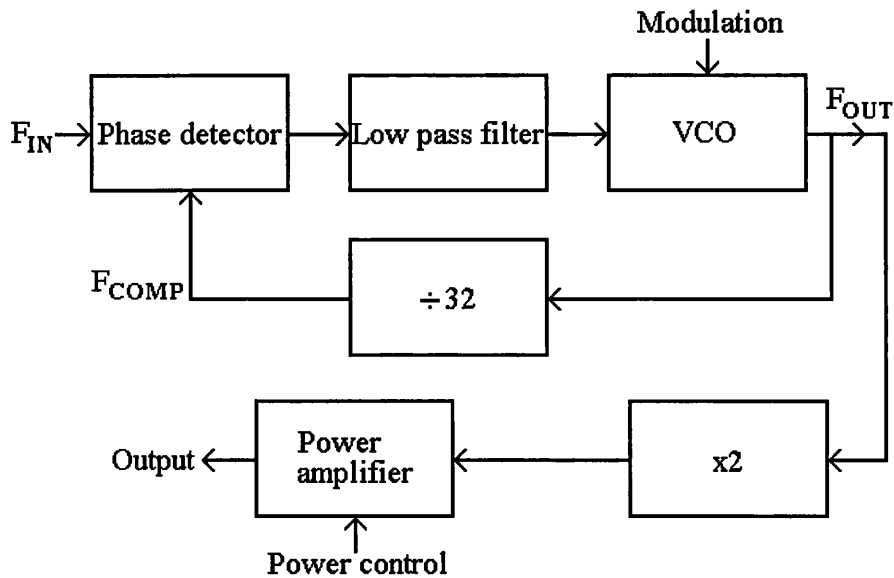
Figure 8.5 -- A block schematic of the radio transmitter

In practice, modulation will be applied to the VCO to provide frequency variations. This modulation is input to the transmitter chip at CMOS signal levels (0 and 5 Volts); it is converted to bipolar form ($\pm V$ Volts) before being applied to the VCO resulting in symmetrical variations about the carrier frequency when the modulation is equally likely to be 'high' or 'low'. If the duration of a voltage level on the modulation input is long enough for the resultant frequency deviation to pass through the phase detector/low pass filter/VCO feedback loop the frequency variation will be "corrected", and little frequency deviation will be present on F_{OUT}. Therefore, during the process of VCO stabilisation no voltage level is applied to the modulation input.

This "correction" of an intended frequency deviation also occurs during data transmission. There is a maximum to the number of consecutive identical bits (that is duration of a modulation voltage level) that can be successfully transmitted. This maximum is determined by the bandwidth of the low pass filter.[12] This bandwidth is decreased by increasing the capacitance such that longer modulation durations are available. However, there is also a need to minimise power consumption by minimising the time needed for VCO stabilisation. This stabilisation time is minimised by *increasing* the bandwidth of the filter by decreasing its capacitance, as

---

[12] which is approximated to the frequency at which there is 3dB amplitude attenuation

higher frequency phase differences will have less aliasing of the phase difference between F_{IN} and F_{COMP}. There is a balance to be struck: it was determined that the maximum duration of a modulation level would be 7.5μs, resulting a maximum VCO stabilisation delay of 3.9ms.

The output from the VCO, including modulation, is then frequency doubled to 916.36MHz. Finally it is passed through a power amplifier, that can be disabled by the application of 0 Volts to the power control input,[13] and transmitted by a quarter wave monopole aerial. During VCO stabilisation this power amplifier is held inactive to minimise transmission duration.

8.3.2 Radio receiver

The radio receiver initially mixes the received signal with a local oscillator to generate an intermediate frequency (IF) of 20MHz, figure 8.6. This IF is then split, with one component passing into a resonant tank that shifts its phase by 90°. The phase shifted signal is then mixed with the original IF, removing the 20MHz component. The resulting baseband data is output at CMOS voltage levels. The process of generating a stable 90° phase-shifted signal in the resonant tank occurs most rapidly if the incoming IF signal is oscillating about the central 20MHz frequency (that is, if the mean frequency deviation is zero). This can be ensured if the initial signal is "1010..." and the duration of the first bit is *half* that of the subsequent bits. When these conditions are applied the radio receiver can reliably output the transmitted data within six bits.[14]

---

[13] the power amplifier raises the radiated power level from -10dBm to +4dBm. This implies that the nearest handset must not be closer than 1m to the receiver if the farthest handset is 22m away (the maximum range for this transmission pair) to ensure that the nearby handset (undergoing transmitter power-up stabilisation) does not interfere with the farther, actively transmitting handset. It is unlikely that in a lecture theatre a student will be within 1m of the lecturer.

[14] in 500 trials the output was found to be stable after these six bits
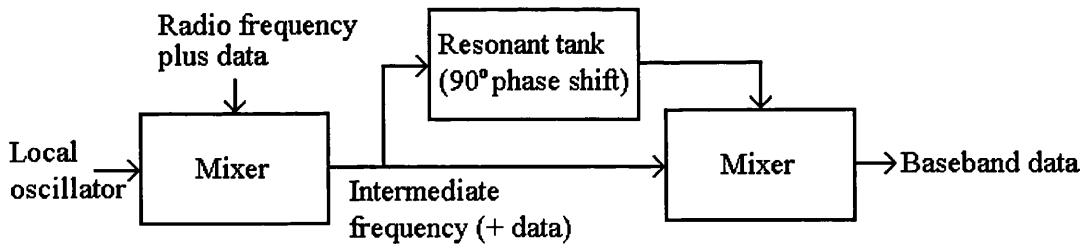
Figure 8.6 -- A block schematic of the radio receiver

8.4 Baseband data recovery

The baseband data output of the radio receiver is at CMOS voltage levels; between transmissions this output is noise. The pulse train for a received transmission is known, and shown in figure 8.7. The first few bits of this pulse train are not accurate reproductions of those transmitted (the bit durations vary from the expected 1μs). As detailed above this output stabilises within 6 received bits. Data restitution is performed using asynchronous clock recovery.
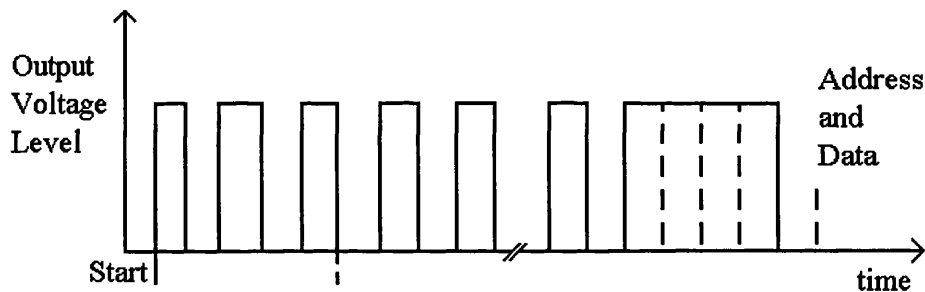
Figure 8.7 -- The baseband output of the downconverted transmission; up to the first six bits have durations that vary from the expected 1μs

In the absence of a signal there is a square wave noise output with amplitudes (0 and 5V) equal to the levels when a signal is present, so it cannot be filtered using simple analogue amplitude or frequency filters. Instead, the output is tested to see whether it has the characteristics of the intended signal using digital processing techniques, figure 8.8a.

The radio receiver output is subject to an initial test: the duration of a received "high" pulse must be within 5% of the expected 1μs. This test rejects much of the noise and

ensures that the subsequent, stricter check is conducted after the receiver has stabilised.

Figure 8.8a -- The baseband output of the radio receiver is subject to an initial check rejecting much of the noise and ensuring the subsequent, stricter test is applied to data that has stabilised. The 32MHz oscillator runs continuously, ensuring stable operation on demand, with its output (shown in purple) not used by the inactive counter

If the output of the receiver passes the initial test a tighter check is applied, figure 8.8b. This tight check function is the application of a comb filter to the data: there must be the correct sequence of edges (rising/falling/rising etc.), and each must occur 1µs (±5%) after the previous. This test is applied to 8 edges (i.e. 4 consecutive bits). If any edge does not occur at the expected time the check is immediately terminated and the initial check once again applied. Once this test is passed the data is adjudged to be a handset transmission and the receiver attempts to synchronise its clock with the handset's clock.

Figure 8.8b -- A tight check is then applied to the data: there should be a sequence of 8 rising and falling edges, each edge separated by 1μs. If any edge is outside this criterion the check is immediately terminated and an initial check reapplied

Receiver/transmission clock synchronisation relies on the fact that the preamble is a series of highs and lows. Once the tight check criterion has been passed the receiver attempts to synchronise its clock with that of the data, figure 8.8c. The data stream is converted into a bipolar form for this function (a low being -5 and a high +5V). An integrator, with an initial value of zero, is activated and integrates this bipolar data over a rolling 3μs period: during the preamble it will not rise above 3.5V. The integrator's output is connected to a comparator with a 5V reference voltage. When the sequence of 4 consecutive 'highs' is received the integrator output is above the reference voltage and the comparator activates a falling edge detector that provides the clock synchronisation.

As a final check for noise, a timer is started at the same time as the integrator and comparator. If the output of the integrator does not become greater than the reference voltage before the maximum possible duration of the preamble this timer

105

deactivates the integrator and comparator, and the output of the receiver is again subjected to the initial check.
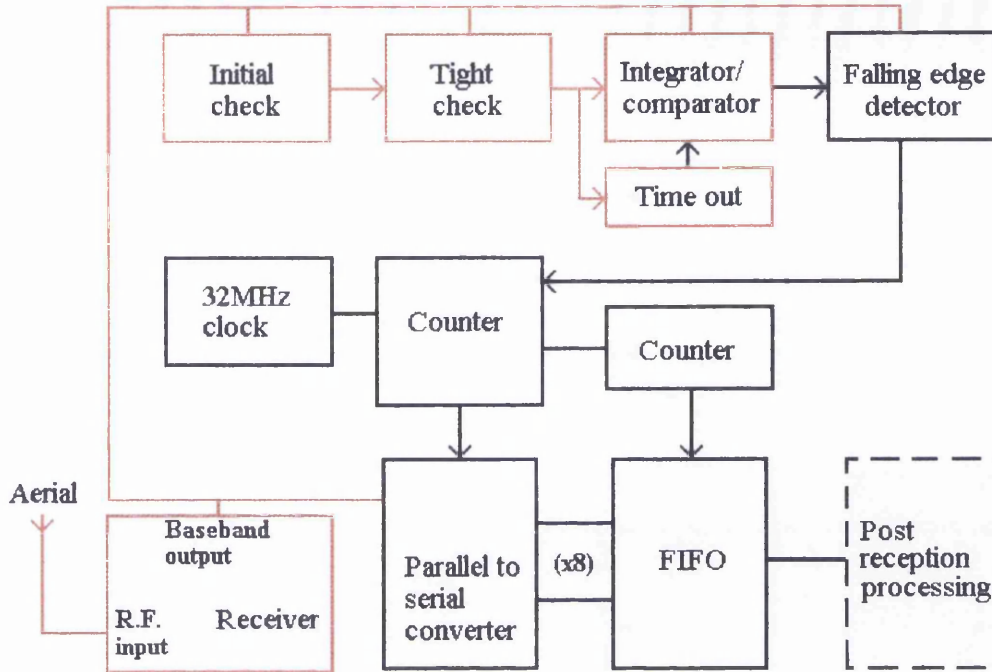
Figure 8.8c -- The receiver now waits for the sequence of 4 'highs' that precede the clock synchronising 'low'

On the reception of the falling edge of the 'low' immediately prior to the commencement of the address and data a counter is started, figure 8.8d. This low acts as the 'mark' bit for clock synchronisation (section 5.1.2). This counter has as its clock input a 32MHz crystal oscillator, whose period is entirely independent of the period of the data. However, the 32MHz oscillator will clock the counter within $0.03\mu s$ of the falling edge of the mark bit - sufficient synchronisation for successful restitution of the data.

106
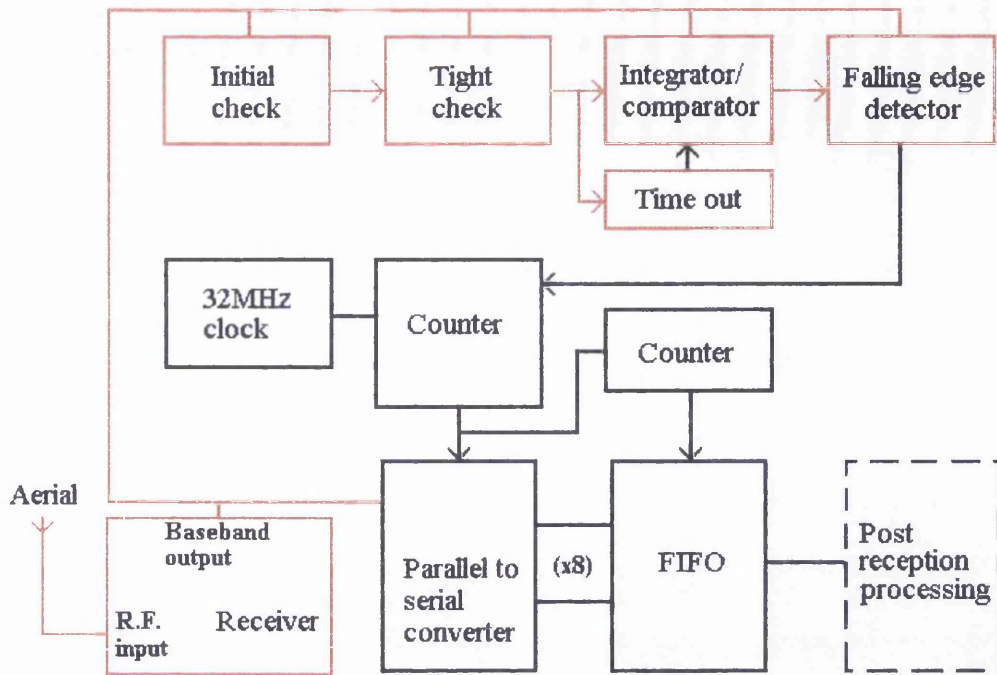
Figure 8.8d -- The falling edge detector activates a counter which has a 32 MHz clock input. This counter will be clocked within 0.03µs; sufficient synchronisation for the 1µs duration of the data pulses

The counter provides a "read" pulse to the parallel to serial converter after 48 cycles of its clock input (i.e. disregarding the first, clock synchronising 'low'), figure 8.8e. This read pulse will occur very close to the centre of the received data pulse, further ensuring that the data will be correctly received (for an erroneous read pulse the clock would have to drift 16 of its total 704 cycles during the address and data recovery - a drift of just over 2 parts in 100!) . Subsequent read pulses then occur every 32 cycles of the crystal oscillator.
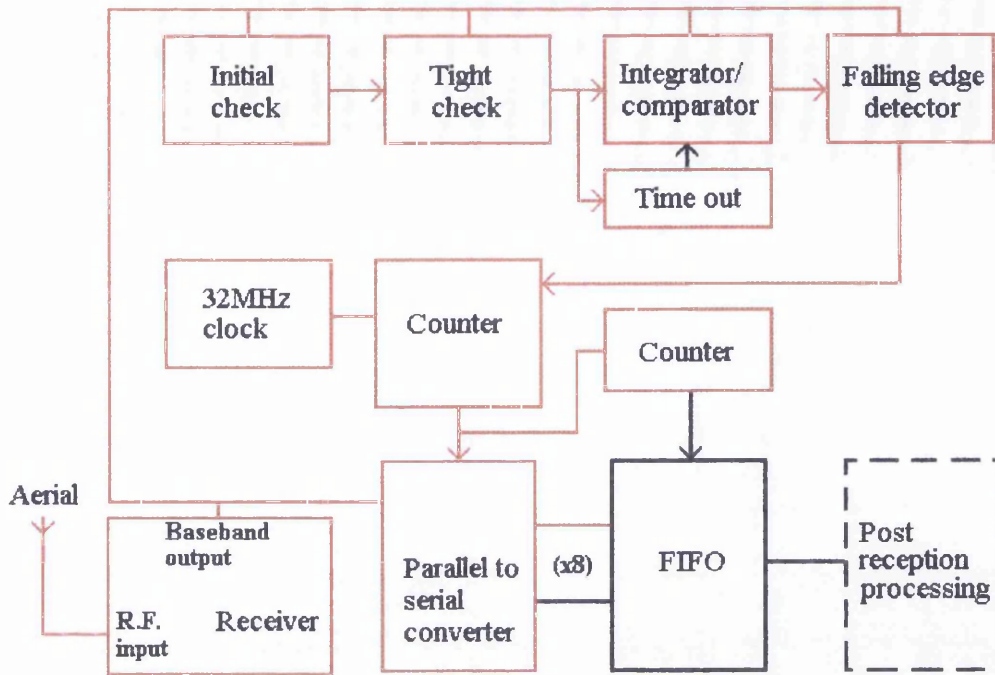
Figure 8.8e -- The counter provides "read pulses" in the centre of each bit of the transmission so that the data is stored in a parallel to serial converter - and these read pulses are themselves counted

A second counter has as its clock input these "read" pulses. After 8 read pulses have been received (the 8 parallel output data lines of the parallel to serial converter consist of the first 8 bits of the transmission), this counter then outputs a "write" pulse to a First-In First-Out (FIFO) memory device, figure 8.8f. After a further 3 read pulses from the 32MHz counter a second write pulse is generated, loading the final 3 bits of transmission into the FIFO. The transmission is now successfully stored. A full circuit diagram of baseband data recovery is included in appendix 5.
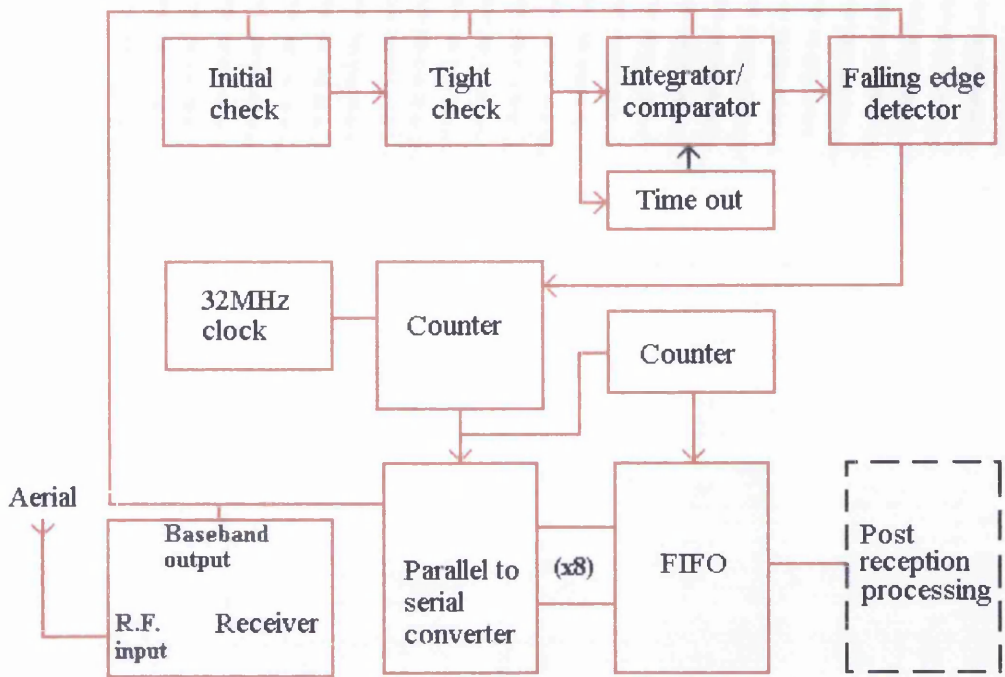
Figure 8.8f -- A second counter provides "write" pulses to a FIFO memory device, in which the transmission is stored. Subsequent post-reception downloading into a computer is carried out in an identical manner to that of the prototype system

8.4.1 Post-reception processing

Subsequent post-reception processing - the download of the stored data into a computer for storage and display - is identical to that of the prototype system (section 7.3). Once again there is a possibility that a collision between transmissions occurs during the process of registration, when the addresses of the handsets present are collated by the computer software and set as 100% of the audience for the purposes of the graphical display, prior to the system's use. This could cause a student's handset to be registered as not present in the lecture; unlike the prototype system, this may not be discernible as the differences in bar height could be as little as 0.5% of the total. If every student pressed their button a very large number of times it is virtually certain that all handsets present would be registered. However, if only a fraction of the system's handsets were in use there is a probability that a bit error in the received address could cause the registration of a handset that was *not* present (a 'ghost' registration) - and this probability rises with the number of times students transmit. Therefore, it is important that the number of transmissions is optimised

such that it is highly probable that all handsets have been registered with a minimised probability of a ghost registration. This problem has been amenable to analysis (appendix 6) and three button presses was determined to be optimal. It is very likely that all students present will have registered their handsets (this probability is greater than 99% under all conditions). The probability of a ghost registration is most acute when around half of the possible complement of handsets are in use (as when there few handsets present it is unlikely that a bit error in the address transmissions will occur, and when many handsets are in use it is less likely that this bit error will create a handset address that is not present). Under worst case conditions, with the probability of bit error maximised and half of the handset complement in use,[15] the probability of no ghost registrations is just under 50%; the probability of one ghost registration or fewer is 85%. As the effect of this level of ghost registration would be to reduce the display height of each student's opinions by a few percent it was determined to be acceptable. It should be noted that when a parity bit is appended to each transmission the probability of a ghost registration falls to virtually zero - if a single bit error occurs in a transmission it is discarded. This parity bit is to be introduced as a matter of priority (section 10.1.3).

8.5 False positives

Although the receiver has been designed to minimise the probability that noise will be stored as data, this probability cannot be entirely eliminated. To gauge the significance of this, the receiver was activated for 12 hours, with no transmissions. When the FIFO was subsequently emptied it had stored a total of 93 "false positives" (data that was not transmitted). This total includes 22 "transmissions" that had button press identification bits that could not have been generated by a handset ("000", "110" or "111") and hence would not have been displayed to the lecturer. A more general expectation would be to have three eighths of the button press identification bits such that they would not be displayed (that is 35 of the 93 received). This level of plausible false positives - on average less than 1 every 10 minutes - can be tolerated as it should not significantly affect the system's use as a

---

[15] all handsets transmitting at maximum range, 20m, with a bit error probability of 5×10^{-4}

means to observe the aggregated opinions of lecture audiences numbering 30 or more.

8.6 Cost per handset

The handsets built as a 'one off' system cost £50 each; this falls to £40 if more than one system is built due to economies of scale (a calculation detailed in full in appendix 7). The receiver will cost less than £100 to construct. This extrapolates to a cost, for the construction of more than one 249 handset system, of £12550. The design lifetime of the feedback system is 5 years; therefore the annualised cost of the system is some £2500 (less than 10% of the cost of one lecturer).

However, if many systems were constructed the cost of each handset would be substantially reduced. All of the components, bar the transmitter, could be placed on an Application Specific Integrated Circuit (ASIC). The chief costs would then be the radio transmitter, box and the buttons. The cost of each handset falls to approximately £15,[16] (a 249 handset system costing less than £4000).

8.7 Handset complement

The modulation may not have more than 7 identical consecutive bits as the modulation voltage level may not be constant for more than 7.5µs (section 8.3.1). This precludes the use of 7 addresses, as detailed in table 8.1; 249 handsets can be accommodated.

---

[16] £7 for the transmitter, £3 for the box, £4 for the buttons, and £1 for the combined costs of the ASIC, handset construction and distribution

Clock synch.	A0	A1	A2	A3	A4	A5	A6	A7	D0	D1
X	0	0	0	0	0	0	0	0	X	X
X	1	1	1	1	1	1	1	1	X	X
0	0	0	0	0	0	0	0	1	X	X
X	1	0	0	0	0	0	0	0	0	X
X	0	1	0	0	0	0	0	0	0	0
X	1	1	0	0	0	0	0	0	0	0
X	0	1	1	1	1	1	1	1	1	X

Table 8.1 -- The transmitter precludes using certain addresses (with bits designated A0 to A7) as the modulation voltage must not stay constant for 8 bits or longer. The preceding clock synchronisation bit is a low; the succeeding data bits may take the value 00, 01 or 10 according to which button has been pressed. Irrelevant bits have been marked "X"

8.8 Probability of transmission failure and misreception

Each handset transmission lasts 40μs. If a full complement of 249 handsets transmit once per second, equation 5.1 (p.36) indicates that the probability of collision is an acceptable 2.0%. This figure indicates the maximum probability of error; a decrease in the number of handsets (and/or the rate of button pressing) will result an almost linear decrease in the probability of collision (for example, for an audience of 62 using their handsets once per second the probability of collision falls to 0.49%).

There is a separate probability of a bit error during the transmission due to the effects of noise. The probability of one or more bit errors in an n bit data word, P_{EX}, where the probability of error in a *single* bit is p_U, is:

$$P_{EX} = 1 - (1 - p_U)^n \qquad \text{Eq. 8.1}$$

If a bit error occurs during the preamble or clock recovery bits it is likely that the transmission would not be recognised by the receiver as the transmission may not meet the checks outlined in section 8.4 or may fail to synchronise the receiver clock. The radio receiver does not recover baseband data accurately during the first few bits of a transmission - the quickest stabilisation observed was 3 bits, the longest 6 bits. The probability of a bit error in any one of the subsequent bits before receiver clock synchronisation, p_U, is 5×10^{-4} at the transmitter's maximum range, 20m. This probability of bit error is used to give the worst case probabilities for the system; quicker stabilisation also results in an increase in the probability of error during preamble and clock recovery (there are more bits in which to have an error). Finally, it is assumed that *any* errors during this period result in transmission failure, to provide a worst case estimation of this probability. Using equation 8.1, the maximum probability of one or more errors during the maximum 26 bits of stable preamble and clock synchronisation is 1.3%.

This probability can be reduced if the signal to noise ratio is increased. The radio transmitter radiates at its maximum power, so to increase this ratio would require a post-transmitter radio frequency power amplifier - not an attractive option in light of the power limitation. This probability of transmission failure was therefore tolerated.

The total maximum probability of transmission loss in a 249 handset system, each transmitting once per second, is 3.3%. There is also a 0.6% probability of one or more undetected bit errors during the 11 bits of data and address at the maximum, 20m, transmission range, equation 8.1; however, some errors will not be displayed by the software as they could not have been generated by a handset (section 8.5).

8.9 Energy consumption per transmission

Each handset transmission dissipates 1.17mJ of energy (the calculation is conducted in full in appendix 8). Commercial non-renewable "AA" size batteries typically contain 15kJ of energy [DURA]; each handset is powered by two. Assuming constant peak handset usage - each handset being operated once per second -

throughout one hour lectures, the batteries would last some 7100 hours. However, a more realistic assumption of the *average* rate of handset use, one press every ten seconds, indicates a battery life of tens of thousands of lecture-hours (that is, the shelf life of the battery expires before it is drained). The batteries do not need regular replacement, and the low-maintenance criterion is satisfied.

8.10 Conclusion

System 2 meets the design criteria for radio frequency lecture feedback. Very large class sizes (up to 249 students) can be accommodated with an acceptable probability of collision, 2.0%, under high use conditions (each of these 249 handsets transmitting once in a second). The handset radio transmissions can be reliably detected at a range of 20 metres, sufficient for use in most lecture theatres. The maximum total probability of transmission failure (including the effects of noise) is 3.3% for a handset transmitting at this maximum range; this handset will also have a maximum 0.6% chance of one or more undetected bit errors in the address or data. The handsets will cost £40 each - indicating a total cost for a 249 handset system of some £12550. The handsets do not require significant maintenance, as the power requirement is such that the batteries' shelf life is likely to expire they are drained.

Chapter 9   Experience with prototype system

"you feel sure that in all the wilderness of possibility, in all the forests of conflicting opinion, there is a vital something that can be known - known and grasped. That we will eventually know it, and convert the whole mystery into a coherent narrative. So that one's true life - the point of everything - will emerge from the mist into a pure light, in total comprehension"   Brian Aldiss, 'Helliconia Winter'

Trials were conducted to gain experience with the feedback system and to gauge the level of its use by students in learning situations. Initially it was introduced during a research seminar day, where the speakers utilised audience prompted feedback (the speed function). It was found that the system was used by the audience throughout an extended period of learning.

The feedback system is primarily intended as a means by which to improve large class lecturing by increasing the quality, quantity and immediacy of student-lecturer communication, with the intention of making lectures more tutorial-like. However, it can also be used as a research tool. The 2 major experiments conducted with the system were a series of lectures to second year undergraduates, firstly to UCL physicists and then to Imperial College electrical engineers. In these talks both lecturer- and student-prompted communication occurred; a total of 54 students attended. Correlations between variables were quantified; there was extensive use of the feedback system throughout these lectures.

9.1   Research methodology

Research into education in the United Kingdom, including higher education, has traditionally followed a paradigm. The researcher seeks to interview the subjects; this interview can be conducted to an agenda via a questionnaire, or in a more free-flowing format in which the subject has greater influence over the interview path. This approach to research has merits, not least that the researcher can hope to probe

the student's reactions in some detail. In turn, this permits the formation of well-argued and reasoned explanations and theories regarding their learning.

However, this methodology involves subjective reaction on the part of the researcher, and it is expensive. The researchers are required to spend time collating and interpreting returned questionnaires, or actually conducting the interview itself. This limits the sample size. One can view the nature of much published research as being *driven* by the tool with which the research is conducted.

There is, however, an alternative choice of methodology. Large quantities of data can be gathered and analysed using automated techniques with a very small requirement for human intervention. This permits the study of samples that are orders of magnitude larger than those gathered by the interview technique. Fortunately, the feedback system permits this via the storage of identified data from handsets (with the time of its transmission). The data can then be analysed en masse by computer and the correlation between variables established. One can also compare handset data with other data external to the lecture, such as previous year's exam mark; or even with items such as where students choose to sit in the lecture theatre. It should be noted that some forms of computer mediated learning also lend themselves to this "event logging" methodology.

With the results of statistical analyses of large samples there is the enticing prospect of applying Popper's principle of falsifiability [POPP]. After observations have been made some numeric confidence in a particular correlation is arrived at. Further experiments are then conducted and one independent variable is quantified. The numeric criterion of correlation is applied to this data and the expected dependent variable calculated. This theoretically derived variable is then compared to reality - and is capable of being proven wrong. If enough tests of the theory are successful (the observed data matching the theoretically derived correlation) the theory can be regarded as sound. This is a new possibility in education research.

Whilst it is well understood that correlation does not imply causality, firmly established correlations can be used to tease it out. If there is no immediate causal link, mediating variables can be sought. With some hypotheses of causal relations event logging could be an appropriate experimental method. However, for some it may be ineffective as it is essentially observational and not explanatory - event logging reveals correlations, and the identification of causal relationships is a very different matter and may require recourse to traditional methods. These methods allow a focussed and in-depth exploration of the discovered correlations. One can ask the students why they reacted as they did to the events that have occurred - *in the sure knowledge that the particular reaction is widespread.* Finally, once causal links have been established one can turn to the central task and aim of education research: improving the learning outcomes of lectures in light of proven results.

There is a further distinct rôle that the feedback system can fulfil: studying the process of the formation of understanding (or misunderstanding) without the intrusion of questioning. The students' opinions on the lecture, and by inference their understanding of it, are directly noted by the speed control function.[1] As previously, this data can then be compared to other variables. Additionally, the lecture can be videotaped. Anomalous features of student-prompted feedback can then be compared to lecture theatre events and the lecturer's actions (with the video recorder clock synchronised to the computer clock). Over a number of similar situations associations can be drawn between these anomalies and events. Once again, this holds out the alluring prospect of being able to apply, statistically, the principle of falsifiability. This can be applied by moving from studying the data and then observing events associated with it to studying the video in the *expectation* of finding an identified associated feedback anomaly (at some probability). If these associations become apparent, there is the tremendously exciting possibility of actually directly improving lecturing methods: the consequences of lecturer actions would be mapped out. Although this research technique mitigates the cost advantages of the feedback system, it is still semi-automated and needs less

---

[1] there is an implicit assumption here that the students will use the feedback system to accurately record their reactions. As the data is confidential to the participants in the lecture there is no reason to believe the students will misrepresent their feelings. In the long term, with the lecturer reacting to their button presses, there is a strong *dis*incentive to misreport: frustration of progress

researcher intervention than the traditional technique - and its potential is so dramatic that it is likely to be a worthwhile pursuit.

At this stage it has been possible to show: that the feedback system is used during an extended period of teaching, there could be identifiable links between lecture events and student-prompted handset use, and correlations between variables are identifiable. These results require further confirmation, but preliminary hypotheses arising from these correlations have been proposed and suitable falsifiable tests for them are suggested, for further research.

9.2 First trial

The system was used during a collaborative research review day, in which there were seven talks on various topics presented, with none of the speakers being involved in this research programme. During all of the talks the speaker was videotaped for later analysis, with a time accuracy between button presses and recorded events of one second. Whilst this is not an identical situation to the target arena, it was felt that it was a reasonable approach to gain initial data regarding handset usage.

The intention was to introduce the handsets into a learning environment and observe whether the fundamental presumptions were supported: that those receiving lectures *did* wish to communicate with the lecturer on their own volition, and that the handsets would remain used throughout an extended period of teaching.

9.2.1 Experimental details

The handsets were taken to a day of talks given between two research groups working in chemical physics and chemistry, held at Reading University. Seven presentations were made by different postgraduate students, none of whom had been previously exposed to the system. All of the participants were informed that the handsets would be used four days before the event itself. At the beginning of the day there was a brief familiarisation talk on how to use the handsets, and an explanation

of the graphical output (including the fact that the student-prompted feedback on the speed of delivery would decay linearly over 30s to prevent the presentation of obsolescent data). There were no further mentions of the system. In the event, no speakers chose to use the question asking facility so the speed function was operative at all times.

9.2.2 Results and discussion

The results show quite clearly that the handsets were utilised throughout the day with pronounced peaks in usage, figure 9.1. It is important to emphasise that these button presses are *purely* audience initiated, and will have some relation to the formative progress of understanding during the talks.

These peaks coincide with events during the talks during the day. The initial peak, in the first 100 seconds, was audience members playing with the handsets they had just been given. The very noticeable peak at 8000 seconds coincided with the speaker apparently becoming confused over the subject material and consisted mainly of 'too slow' with approximately 10% 'too fast'. Subsequent questioning of the speaker confirmed this observation as he reported "I got confused about what I was saying". The data peak during the delay, shortly before 12000 seconds, was almost entirely 'too slow'. The break on the abscissa was lunch.

Whilst it is possible to exactly associate the level of handset activity with talk events, with this quantity of data it would be an unrewarding exercise. The scrutiny of the responses of 8 individuals will not shed light on a more generalised audience with its inevitably more variable nature.
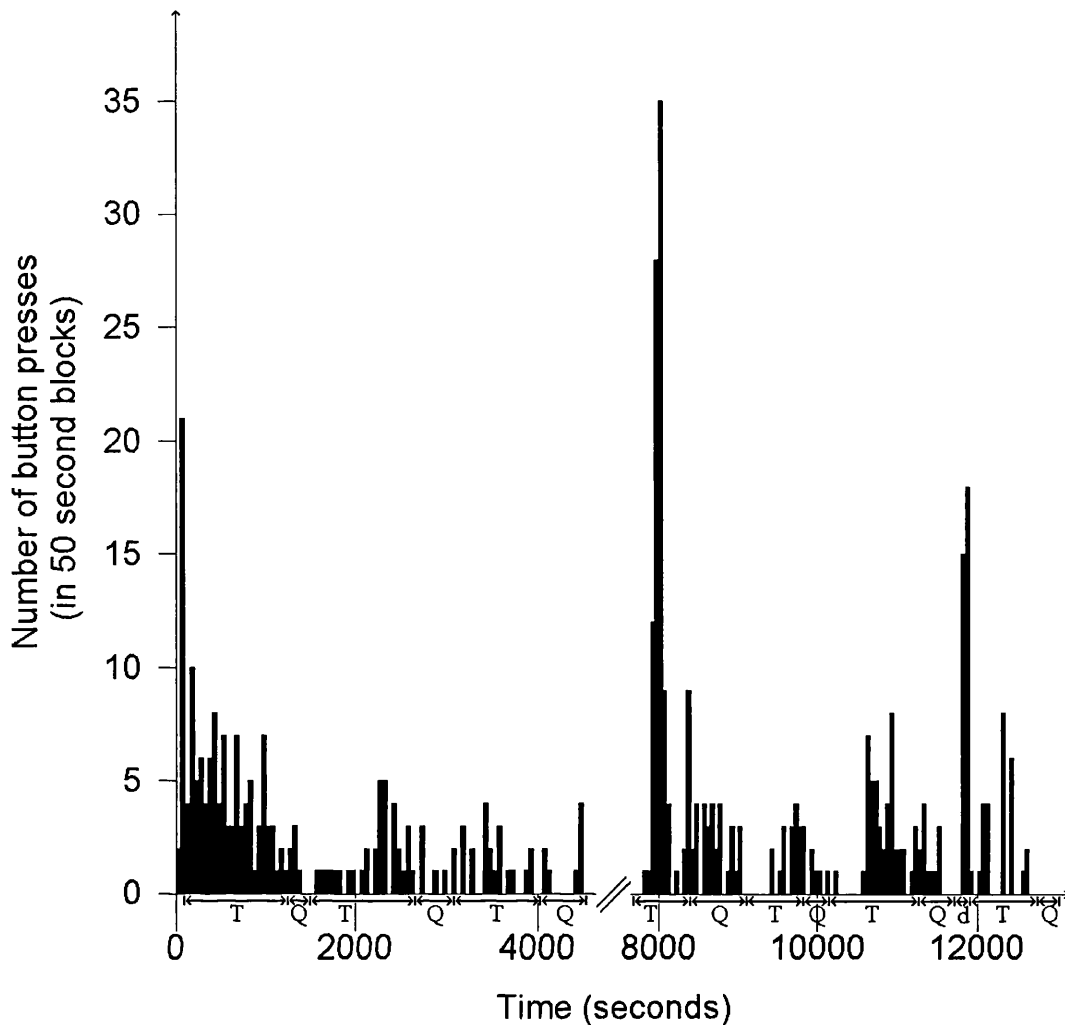
Figure 9.1 -- Self-prompted feedback from 8 handsets during a day consisting of a series of talks, T, and question sessions after, Q. The lunch break is indicated; 'd' was an unplanned delay. The data has been collated into 50 second blocks.

9.2.3 Conclusions

This data demonstrates that the feedback path is robust: handsets remained in use throughout an extended period of learning. It also demonstrates that it is possible to analyse handset activity and relate it to lecture events as posited earlier.

On a practical level, it became obvious during the day that the decay time (30s) for the speed function was too rapid - the frequency at which a speaker would glance at

the screen was overestimated. After analysing the video and data, it is now felt a more appropriate decay time would be one minute.

9.3 Covariance trials

After the successful trial of the feedback system at the collaborative research day, confirmation that undergraduate students would appreciate the feedback path was sought. Additionally, the relationships between several variables were recorded and compared: students own opinion of their understanding during its formation, the proportion of the questions posed during the lecture answered correctly and the student's previous year's overall mark. For these purposes a series of small-class talks to a total of 54 second year students were given.

9.3.1 Experimental details

The author gave the talks to second year students to a sample of 22 students from University College London Physics department and, in light of the surprising results, a further sample of 32 students from Imperial College Electronic and Electrical Engineering department.

It was felt to be of critical importance that the talks should be on a topic unfamiliar to the students, to ensure that the system would be used in as realistic a learning environment as possible. Students were compelled to attend to ensure that the samples were *not* self-selected. The subject of the talks to the physics students was basic information theory,[2] and the subsequent talks to the electrical engineering students were regarding atmospheric spectroscopy.[3] The overhead projections used during these talks, including the questions posed to the students, are included at appendix 9.

---

[2] topics covered included Shannon's Law and the vector representation of modulation
[3] covering the atmospheric dynamic energy balance, the greenhouse effect and the consequences of atmospheric ozone depletion

The maximum class size necessitated by the prototype feedback system used, eight, represents a very small lecture. The author attempted to mirror traditional large-class lecturing by ensuring the talks had an absolute minimum of (verbal) feedback from the audience.

These talks differed from mainstream tertiary education lectures for three fundamental reasons:

- the students knew that they were not going to be subsequently examined on the material introduced
- the class size used represents a very small lecture
- the use of the feedback system itself

These differences prevent entire confidence in the tests being applicable to mainstream lectures. However, there were similarities with lectures:

- the teaching itself comprised of monologues with little or no verbal feedback from the students during the talk itself
- substantially new material was introduced to the students
- attendance was compulsory - the sample was not self-selected
- other than the feedback system, only overhead projection was utilised

In addition, there is no *a priori* reason to believe that that the feedback system (designed to used by students with minimal effort or distraction) would influence the nature of students' learning - with the exception of any Hawthorne-like effect.[4] On balance, I believe these lectures were not substantially dissimilar to those ordinarily given to undergraduate students.

During the talks the data was stored with the handset address and time of reception. A mark was stored on the hard disk file when the author switched functions to pose a multiple choice question, allowing differentiation between answers to multiple choice questions and opinions on the speed of delivery. The student's previous

---

[4] the Hawthorne effect is where the introduction of new technology *itself* may cause students to react positively to, and hence succeed in, an educational situation. First presented in [ROES], the nature and extent of the Hawthorne effect has been subject to considerable debate [ADAI][PARS]

year's overall university exam mark was also subsequently noted. These data were collated.

A number of assumptions were made regarding the data:

- the percentage of questions answered correctly by the students during the talk is a reasonably close measure of their level of understanding
- the transmitted indications on the students' reaction to the speed of delivery is a representation of their true feelings
- the previous year's overall mark provides a workable approximation to their total summative learning in the preceding academic year

The anonymity of handset use and the fact that the display was only shown to the lecturer supports the first two presumptions, as students should have felt a minimum of peer pressure and fear of appearing ignorant; the questions were also quite carefully designed to require reflection by the students on the subject matter (and not simply factual recall). Supplementing this, students were reassured explicitly at the start of the talks that results would not be associated with individuals. The third presumption is simply that made in the university system as a whole.

In addition to the data provided by the handsets and exam results, the students were verbally and informally questioned about their reaction to the feedback system at the end of each talk.

9.3.1.1 Lecturer-prompted feedback

During the talks a multiple choice question was posed, after a discourse on a topic, to test the student's understanding. Each talk lasted approximately 30 minutes and 8 questions were asked. The questions were intended to stimulate the students' interest and test their understanding by requiring a degree of thought; questions that purely asked for factual recall were avoided. For example, during the talks on atmospheric spectroscopy there was a discourse on the reflection and absorption of incoming solar radiation by various atmospheric components, amongst which the idea was

introduced that shorter wavelength light is scattered preferentially. After this students were asked the following question:

Q. "Why is the cloudless sky blue?"

    A. Oxygen (O_2) is slightly blue

    B. The sky isn't blue - this is an optical illusion

    C. Because of differences in the scattering of light

    D. Because clouds leave a blue residue

    E. Because it's a reflection of the sea.

This question required: the basic knowledge that blue is short wavelength light (reasonable for students having a physical science secondary education), the ability to recall the Rayleigh scattering mentioned during the talk, and the association of these two ideas. This was a question of typical difficulty. All questions had five potential answers: the correct answer, a very plausible incorrect answer, a less believable answer, a rather implausible answer and one fantastic answer (in the example above these correspond C, A, B, E and D).

9.3.1.2 Student-prompted feedback

Students could initiate feedback by expressing opinions on the lecturer's speed of delivery whilst teaching was occurring. The author found it relatively simple to monitor this - a glance at the screen sufficed to get a good impression of the students' thoughts. On several occasions, material was restated when a significant proportion of students felt the delivery too fast (it was rare to find students viewing the speed of delivery as too slow). Reiteration occurred when the on-screen data indicated above approximately forty percent 'too fast'.

9.3.2 Results and discussion

In the analysis of these results, it is presumed that these two series of talks can be treated as two individual *contemporaneous* trials - that is, it is presupposed that the author succeeded in making every University College talk substantially similar in content and delivery; and, in a separate trial, every Imperial College talk also

substantially similar. Consequently, the following discussion aggregates the University College data into a single analysis and the Imperial College data into a separate, single analysis. This is intended to be a reasonable first approximation.

Correlations have been sought throughout the discussion on the numerical data. To attempt to gauge the significance of those identified the product-moment correlation coefficient has been determined (the data and calculation is shown in appendix 10). This represents the observed data as a normally frequency-distributed, fair sample of a theoretical population. A significance test is then conducted with respect to a theoretical larger population with the sample, against the null hypothesis that there is *no* correlation. This permits the estimation of the likelihood that a correlation observed in the sample exists in the wider student population.

The correlation coefficient determines whether there is a linear relation between the two variables. The correlation coefficient will be 1 for perfectly aligned data on a line with positive gradient, -1 for perfectly aligned data on a negative gradient, and 0 for data exhibiting no correlation. The modulus of the sample's product-moment correlation coefficient can be compared to values in 'look up' tables at 90%, 95%, 99% and 99.9% significance levels (these levels vary according to sample size) [SIEG]. If the observed correlation coefficient exceeds the 90% level one can have some confidence that the correlation is significant; exceeding the 95% level indicates high confidence in the correlation; 99% very high and 99.9% extremely high confidence.

9.3.2.1 Informal questioning

No students exhibited an overtly negative reaction when asked their opinion of the system. Typical responses include: "It's fun - I'd like it in some lectures I have now!" and "I like the questions. They keep you interested trying to work them out". The least enthusiastic response was: "It was all right". Many students showed unprompted interest during this part of the talks regarding how the system would and could work in mainstream lectures. This is encouraging and, whilst by no means a

stringent test, lends credence to the idea that students would appreciate use of the feedback system in lectures.

9.3.2.2 Engagement

The frequency that students used self-prompted feedback system and their understanding during the talks, as manifested by the number of multiple choice questions answered correctly were tested using the product-moment correlation coefficient and found to have exceeded the 95% confidence limit in both trials. The Imperial College data has a correlation coefficient of 0.43 with a 95% confidence level of 0.35; the University College data has a coefficient of 0.44 with a 95% confidence level of 0.42. This indicates high confidence in this correlation. The data, together with lines showing the least squares fit linear regression to the data to provide *indications* of the correlations, are shown in the scatter diagrams of figure 9.2.

This result does not show that student initiated handset use *of itself* increases the proportion of questions answered correctly. Students who are cognitively engaged in the talk are likely to show increased handset use: in the extreme, one must be listening to hold an opinion on the speed of delivery. However, it does seem to provide endorsement of the idea that as students' involvement increases there is a parallel increase in their comprehension. The feedback system attempts to engage the majority of students by the posing of multiple choice questions, and hence giving them a more active rôle in lectures.

A further possibility is that those students who used the speed function frequently had misunderstandings cleared: on several occasions during talks the author restated material when a significant proportion of students (about forty percent) were indicating dissatisfaction. However, this possibility is not supported by later evidence comparing the level of dissatisfaction expressed by students with the evidence of their immediately summative understanding (section 9.3.2.4).
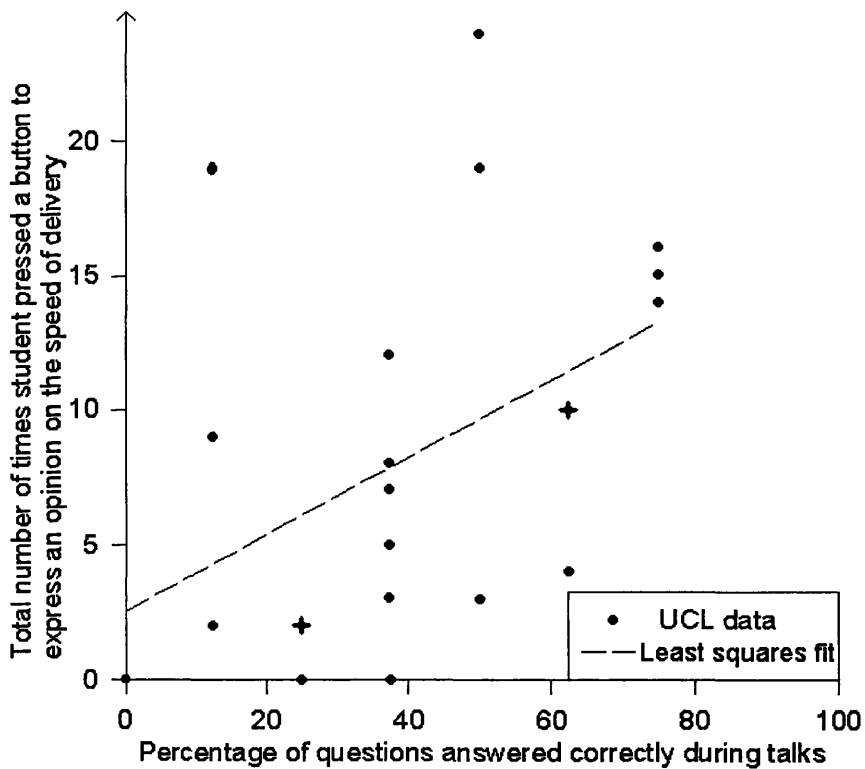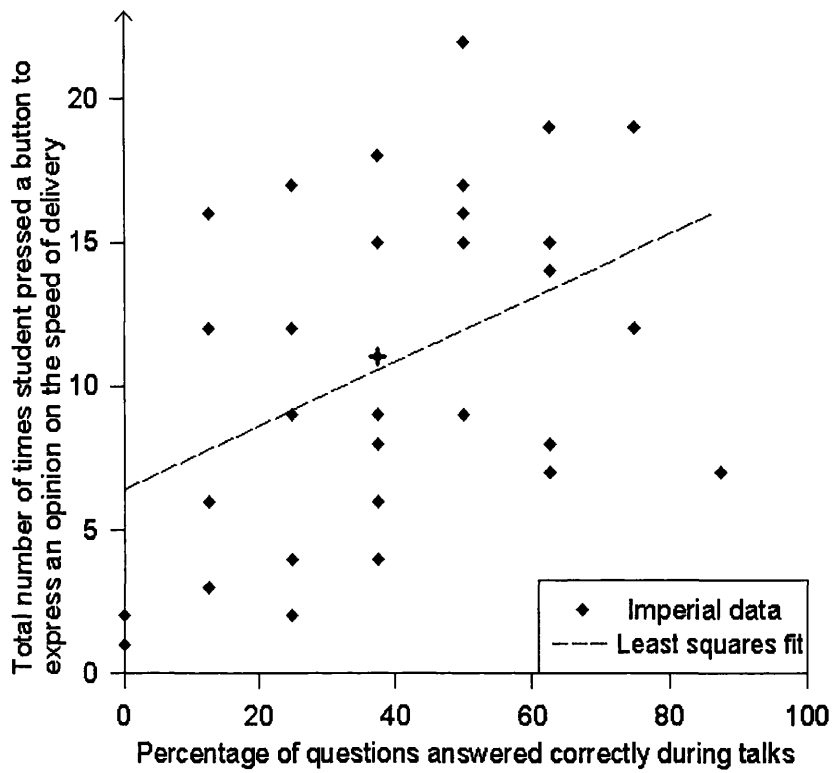
Figure 9.2 -- Evidence of understanding versus student initiated handset use; crosses mark two exactly coinciding data points

One means by which it may be possible to determine whether use of the speed function directly correlates with an increase in the percentage of questions answered correctly would be to hold a series of lectures with multiple choice questions, but with half of them having the speed function unused. It would then be possible to track individual student's answers to questions during both types of lecture. From this, it will be possible to see if high-scoring students are also heavy users of self-prompted feedback - that is, the feedback is simply reflecting an already high level of cognitive engagement from these students in the lecture. It will also be possible to observe any effects of the use of self-prompted feedback on all of the students by comparing the proportions of questions answered correctly in both types of lecture. If there is a general increase when the speed function is in use then it can be concluded that this form of feedback has a positive effect on the observed understanding.

9.3.2.3 Aptitude

Far closer correlations were found between the students' overall mark from the previous year and the proportion of questions they answered correctly during the talks. These correlations are significant to the extremely high confidence level of 99.9%. The Imperial College data has a correlation coefficient of 0.80 with a 99.9% confidence level of 0.56; the University College data has a coefficient of 0.74 with a 99.9% confidence level of 0.65. Figure 9.3 shows scatter plots of these results, together with an indicative line showing the least squares fit linear regression to the data. These correlations are surprisingly close for both sets of data, and strongly imply a link.
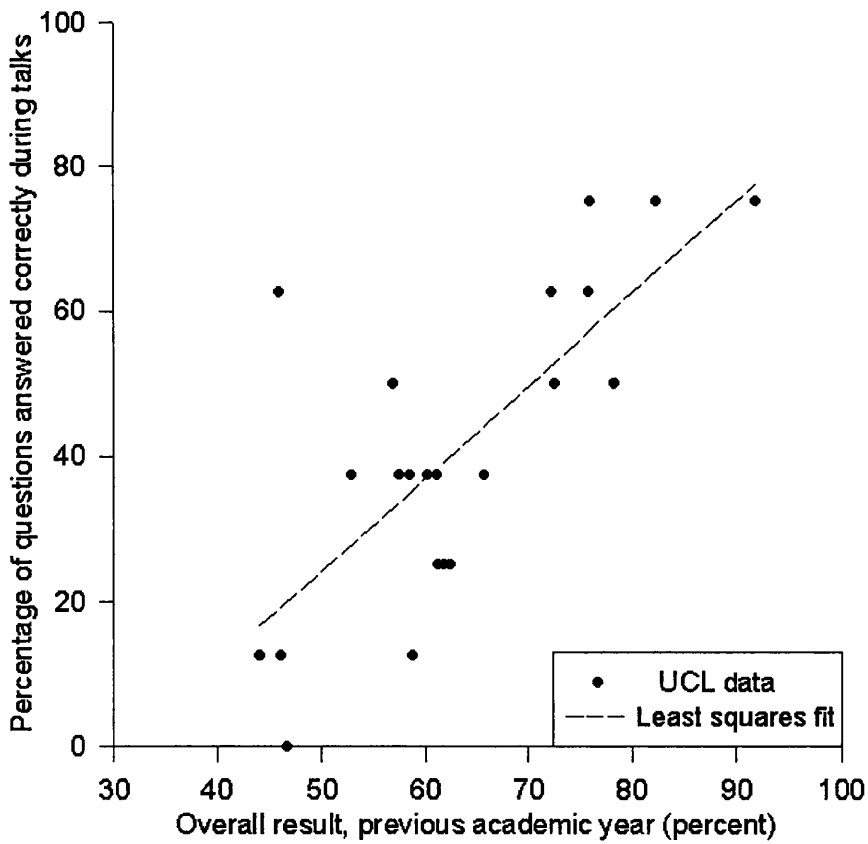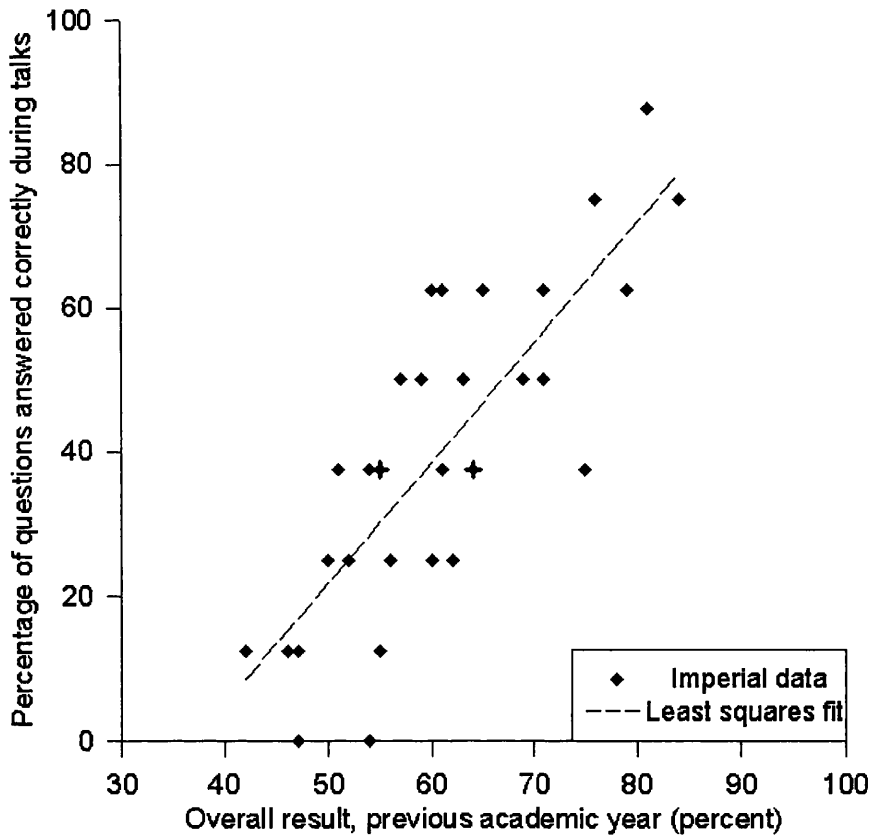
Figure 9.3 -- Previous year's overall result versus percentage of questions asked during the lecture that were answered correctly; crosses mark 2 coinciding points

The proportion of questions answered correctly provides some insight as to the students' immediately summative understanding during the talk. The previous year's overall result provides evidence of the students' summative understanding in a variety of subjects and time scales over an academic year, and comes from many modes of learning (lectures, private study, student-student teaching, etc.). Whilst correlation does not necessarily indicate causality it is reasonable to propose that it *might* do so. If so, this would mean these two items may be linked by some variable or variables on the part of the student. These variables can perhaps be denoted as 'aptitude', with students obtaining higher marks when they have higher 'aptitude'. These putative variables could involve some or all of three properties of students:

- capability in associating the ideas presented to them during questioning with ideas introduced during teaching
- recall of data from teaching
- comprehension of concepts and data during teaching leading to a greater proportion of questions being answered correctly

If the same students could be tracked in several different lectures with multiple choice questioning (and preferably entire lecture courses), wider measurement of the proportion of questions answered correctly would be possible. This would enable an assessment to be made of the variability in an individual student's correct response percentage - the variability of their aptitude. Further, if students used the system across a variety of courses it would be possible to compare their aptitude in different topics. If consistently high, this would indicate that aptitude *is* an inherent, conserved item in students' learning.

Data recollection could be one of the variables in the identified "aptitude". To explore this, it would be enlightening to ask multiple choice questions during a lecture that only called for factual recall in addition to questions that require more reflection on the part of the student. The percentage of factual recall questioning that students answered correctly could be compared to the percentage of more reflective questions answered correctly.

9.3.2.4 Self-satisfaction

Students were able to indicate the level of satisfaction they felt with the state of their learning during the talks by use of the speed function - use of "too fast" or "too slow" showing dissatisfaction, and "about right" satisfaction. Comparison of this to the evidence of their learning as gauged by the percentage of questions answered correctly reveals only some evidence of correlation. The possible correlation has a negative gradient in this sample; the University College data has a coefficient of - 0.53, exceeding the 95% confidence level (where the modulus of the coefficient is 0.46). However, the Imperial College data does not exhibit correlation, with a coefficient of -0.24. The 90% confidence level is where the modulus of the coefficient is 0.46. This indicates that one can be reasonably sure that there is *no* correlation in this data. The scatter diagram of both samples is shown in figure 9.4.[5]

This seems to indicate that students' own perception of their learning may not be a good indicator of their understanding. Students cannot be relied upon to accurately apprehend their comprehension - a surprising result, as one would have thought that they would be best placed to judge it.

It is interesting to note that the students also were relatively pessimistic with regards to their understanding, with few indicating that they were more often happy than unhappy with their comprehension. Students may only use the system when they are discontent - pressing a button when dissatisfied. This could reflect a general mindset on the part of students, one where *inaction* most often indicates satisfaction. This would invalidate the assumption that the transmitted data on the speed of delivery was an accurate reflection of the students' true reaction.

---

[5] it was not possible to ascertain the satisfaction of three students in the University College data as they did not use the speed function at all
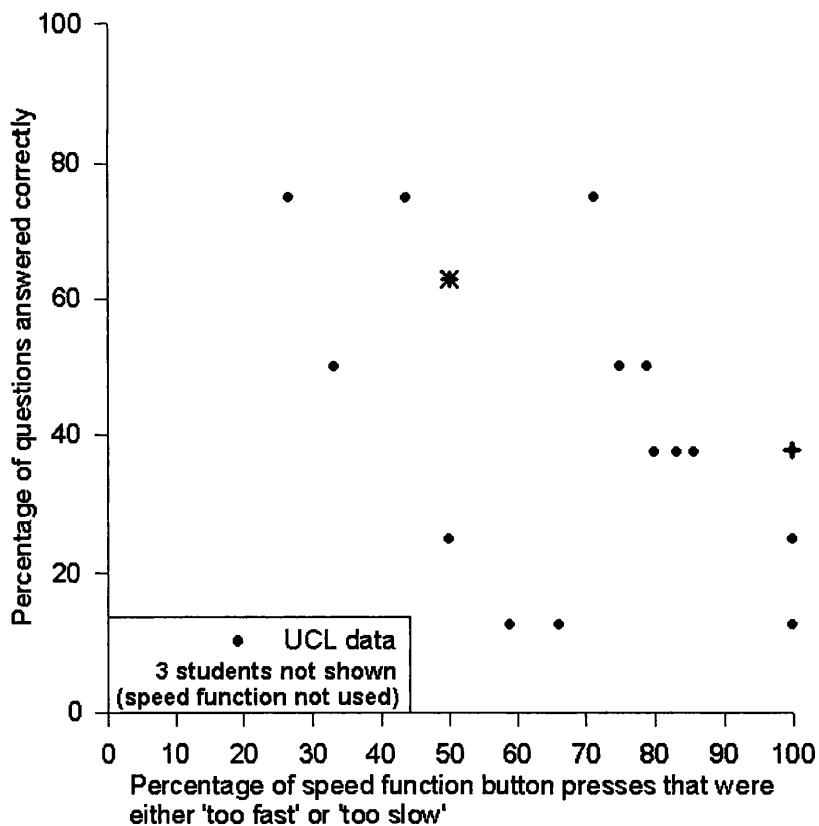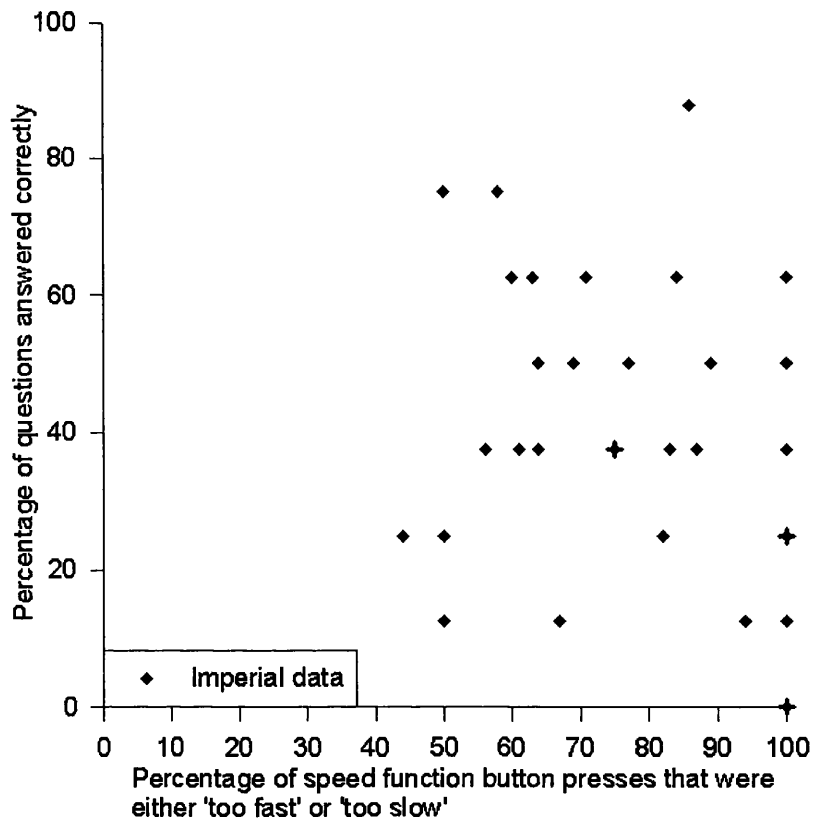
Figure 9.4 -- Percentage of self-prompted feedback indicating student dissatisfaction

versus percentage of questions asked during the lecture that were answered

correctly. Crosses mark 2, and stars 3, coinciding points

There is another, more remote possibility. Students attending the talks had never, to the author's knowledge, been asked to judge their understanding *whilst* the process of teaching is underway (in contrast to end-of-discourse questions, for example). Self-judgement of understanding could be a skill; if so, it is one in which they probably have little proficiency. This leads to the falsifiable hypothesis that their accuracy will rise with time as they became used to, and more expert in, their self-assessment of their learning as they influence the lecturer. This can be tested by simply allowing the students to use the system with a responsive lecturer over a lecture course. If this hypothesis were the case then the system would be providing a more valuable service than had been anticipated: it would be encouraging the students to better understand their *own* learning - a vital asset in a 'learning society'.

9.4 Conclusions

These talks have provided a confirmation of students' use of feedback during a lecture-like situation. No difficulties encountered by students were felt to be serious enough to mention to the author during informal verbal questioning after the talk.

The precise quantitative nature of the correlations in the University and Imperial College data are unlikely to be significant: the topics taught to each were different, and a Hawthorne-like effect may act to increase the absolute activity of the students with this brand new feedback technique. This has no consequence with respect to their qualitative existence.

The correlation between handset use and the evidence of immediately summative comprehension is statistically significant. This could be due to a naturally increased level of activity by those students with strong cognitive engagement in the lecture. Further confirmation and tests of falsifiable hypotheses are needed to bring this result into sharper focus.

Immediately summative understanding and the evidence of summative understanding over much longer time scales exhibit a very strong correlation. In the first instance it

is presumed that there is a mediating variable, 'aptitude', between these two variables. More widespread trials are needed to both confirm the correlation and provide data against which to test hypotheses.

The doubt regarding a correlation between students' own perception of their understanding during its formation and that shown when questioned during the lecture is surprising and inconclusive. It needs to be further explored by more trials. There is a possibility that judging one's own understanding is a skill to be developed which could provide an exciting unforeseen effect from the use of feedback as this skill would be helpful to lifelong learning.

The prime need for future research is for the expansion of the trials to larger class sizes and into a variety of formal lecture courses to enable further validation of these preliminary results, both to determine how widespread the identified correlations are, and, if confirmed, to permit greater exploration and examination of possible causes.

Chapter 10 Conclusions and future research directions

"the world moves, and it moves both very swift and very slow. Swift, because they themselves change little and all else fleets by: it is a grief to them. Slow, because they do not count the running years, not for themselves. The passing seasons are but ripples ever repeated in the long long stream. Yet beneath the Sun all things must wear to an end at last" J. R. R. Tolkien, 'The Lord of the Rings'

"Don't take life too seriously. You'll never get out alive" Bugs Bunny

This thesis addresses a key problem in the delivery of education in universities. The nature of this problem is the disjuncture between students and the lecturer as has been described in chapter 3. The thesis is concerned with a technological solution involving radio links between students and the lecturer with the data being displayed by a computer. The system to meet the objectives has been designed, tested and then further developed to meet the requirement of large classes. It has performed well; it has been designed to be affordable. The system has been tested in three universities; whilst experiments with larger classes are needed, the preliminary results are highly encouraging.

During the development of the work described in this thesis it was appreciated that there was a completely different and exciting outcome - the realisation that the system was extremely well adapted for the elucidation of a variety of educational issues. It permits direct *measurement* of educational behaviour in a lecture without disturbing the lecture itself. Preliminary results indicating the value of this research methodology have been obtained.

There are two distinct areas for future research: technological improvement (including cost reduction) of the feedback system and the research into lecturing that can be conducted with it.

10.1 Technological research directions

The system we have designed and tested works well and meets the original design criteria. However, the large scale usage which is advocated would benefit from technological developments in a number of different directions. It would certainly be possible to utilise the system for very large class sizes (i.e. greater than 249), and indeed simultaneously for a local and remote audience. The following section includes brief indications of some "next-step" developments, and some additional strategic changes that are envisaged.

10.1.1 Cost

The cost of the handset will be reduced as the volume of production increases - if every university bought two sets of 250 that is 50,000 for the UK alone. All of the circuitry, bar the radio transmitter, can be placed upon an application specific integrated circuit (ASIC). The principal costs would then be the transmitter, box and buttons; one could confidently anticipate a reduction in the cost to £15 per handset - £3 per year over the system's 5 year design lifetime.[1]

10.1.2 Transmission failure improvement

The principal current cause of transmission failure is collision between transmissions. This probability can be mitigated by reducing the duration of each transmission. The modulation is currently conducted at the maximum guaranteed rate of the transmitter, 1Mb s^{-1} [RF99t]. It has become evident that this rate has the potential to be significantly increased without much change and without affecting the robustness of the transmission link, to perhaps 2Mb s^{-1}. This would almost halve the probability of collision between transmissions (equation 5.1, p.36). Clearly, one can have transmitter/receiver combinations that permit modulation frequencies one or two orders of magnitude higher. However, the need to use inexpensive components has so far prevented the system from going to such higher modulation speeds.

---

[1] £7 for the transmitter, £3 for the box, £4 for the buttons, and £1 for the combined costs of the ASIC, handset construction and distribution

The probability of an undetected bit error after clock capture, during the transmission of the button press data and handset address, can be reduced by utilising coding - a parity bit permitting error detection (section 6.3.3) and block codes allowing error correction (section 6.3.4). These code bits would be in *addition* to the data and address bits as the system modulates the data at the maximum possible speed: the transmission duration and energy consumption of the handset would increase with no reduction in the energy per transmitted bit.

The addition of a parity bit would decrease the probability of an undetected bit error in the data plus address significantly (section 6.3.5). The probability of collision is now low enough for the introduction of a parity bit to be profitable - with the current modulation speed the addition of a parity bit increases the probability of collision by 0.05% but allows the system to detect those transmissions with a bit error. The additional energy cost is minimal.[2] Erroneously received transmissions can then be disregarded as it is not critical that every transmission is displayed.

The introduction of error *correcting* codes will be attractive when a lower probability of collision has been achieved. The correction of a single bit necessitates the addition of 3 code bits (section 6.3.4), increasing the probability of collision between transmissions. This probability falls to a level comparable with the probability of a bit error, 0.6%, when the transmitter modulates at some $3.3Mb\ s^{-1}$. If 3 code bits replaced the parity bit in transmissions at this modulation speed the probability of collision only marginally rises, by 0.03% (equation 5.1), but data that was previously discarded can be recovered. Although correct restitution of every transmission is not critical, it is desirable.

---

[2] there will be at most a 0.03% increase in energy consumption per transmission (appendix 8)

10.1.4 Spread spectrum

The simplest possible system was chosen largely because of the power limitation in the handset. Nonetheless, the advantages of using the spread spectrum modulation techniques have been evident throughout this research, as this would permit a substantial decrease in the probability that transmissions will interfere with each other (section 5.4) and provide protection against narrowband noise (section 6.2). In particular, Direct Sequence Code Division Multiple Access (DS-CDMA, section 5.4.4) is particularly alluring, as it requires minimised extra complexity in the handset, with the complexity concentrated in the receiver which could use interference cancellation techniques.

To achieve sufficiently low cross-correlation between handset transmissions requires many spreading code modulation bits to be applied to each transmitted bit to spread the transmitter's spectrum sufficiently. The total duration of transmission, and hence handset activation time, must be short enough to meet the handset power limitation. This implies that the transmitter used must have a high modulation rate. To estimate the minimum viable rate, it can assumed that the transmission would include 3 (data rate) preamble bits. The subsequent transmission will not use codes, with the process of synchronisation between the receiver and transmission pseudonoise modulation being used to recover the data clock (section 5.4.5, [DIXO pp.214-260]). Ten spreading bits per data/address bit will be applied [FANF]. This indicates that a total of 140 bits need to be modulated onto the radio frequency carrier. The duration of the handset transmission will remain constant at the current 40μs (to hold the energy consumption approximately constant). This implies a transmitter modulation frequency of at least 3.5Mb s^{-1}. This data rate is not greatly above that currently achievable (section 10.1.2). Therefore, in the future I would expect this to be a profitable avenue for technical research to improve the system. The CMOS output of the radio receiver can be passed directly into a computer and the post-reception signal processing can be performed by software (including 'digital signal processing' techniques [PORA][MULG]). This would enable the signal compression process to be conducted at very high speeds in comparison to the transmitter's data rate (in the

order of at least hundreds of MHz compared to tens of Hz). This would permit the introduction of DS-CDMA without expensive bespoke hardware.

10.1.5 Power

In the longer term, when the energy consumption per transmission falls below $500\mu J$, the use of a photovoltaic power source becomes possible - which is appealing as it would all but eliminate the maintenance requirement with no need to change batteries at all.

The current energy requirement is $1.17mJ$ per transmission. This energy requirement would be significantly reduced if the time required for the transmitter phase lock loop (PLL) to lock could be reduced. This time is dependent on the loop filter, appendix 4, which is simply a capacitor and resistor connected in series to ground. The PLL will lock most rapidly if this capacitance is minimised; however, modulation requires a larger capacitance (section 8.3.1). These demands can be better reconciled if the loop filter were a resistor and *varactor*. The voltage on the varactor could initially be set such that the capacitance is minimised, and could then be varied so that the capacitance is increased to levels where modulation can take place. This variation in the voltage is easy to achieve by, for example, using the voltage charging curve of another capacitor and resistor connected to ground. Initial experiments confirm that this is possible and may offer up to a 70% reduction in PLL lock time - an energy consumption per transmission reduction of at least $300\mu J$.

The introduction of extra features in the handset to provide more options for electronic lecture theatre communication (such as a small screen in the handset to provide direct lecturer-student communication) is not envisaged until the need is proven by considerable experience with feedback. Educational technology should only address *specific* weaknesses in a teaching situation (chapter 3). Until a problem has been identified there is no requirement for a technological solution.

10.2 Educational research directions

The original, and principal, purpose of the system was to allow students a means by which to communicate with the lecturer to permit large class lectures to adopt more of the characteristics of small class teaching: a two-way discussion of the topics being studied. The ability to use the system as a research tool was subsequently recognised. This research can provide a better understanding of the lecturing process; once it has been characterised improvements can be made to it. The following sections detail studies in which the system can be used to improve the discernment of features of the lecturing process.

10.2.1 Confirmation of preliminary results

The experiments at University College London and Imperial College of Science, Technology and Medicine have revealed initial evidence regarding the existence of correlations between the proportion of lecture questions answered correctly by students, their previous year's overall result and the level of satisfaction indicated by their use of self-prompted feedback. The preliminary experiment and the correlations observed are strongly suggestive - but will require further assessment with substantially larger sample sizes.

10.2.2 Lecture theatre geography

The use of the system with such large sample sizes would also permit an exploration of other features of lecturing. These include whether the commonly held belief that lecture theatres have 'geography' is correct - whether able students sit at the front of classes with a decline in students' ability (and/or attentiveness) with distance from lecturer. This could be simply done by placing 'tear away' stickers on the front of each handset marked "front", "middle" or "back"; students can be asked to remove the stickers and so provide this information. An alternative method would be to place the handsets in a known order and videotape the students as they collect them. This permits direct and minimal impact observation of this datum. One can then

study the rate of student prompted feedback and the proportion of lecture theatre questions answered correctly with respect to student' location within the theatre.

10.2.3 Lectures: a single teaching experience?

There is a dynamic within any lecture: the lecturer is providing one teaching experience, whilst the students may be undergoing many different learning experiences - which may or may not be contemporaneous, or even similar. The dynamic between the two can be explored by examining the record provided by the student-prompted transmissions. As these transmissions are individual and confidential there will be little chance of collusion between students affecting the results. If the single teaching experience dominates the expectation would be to find a high proportion of students indicating dissatisfaction contemporaneously and in the same direction (fast/slow); conversely, if the student learning experiences dominate one would expect to find less pronounced peaks. The extended use of the system at Reading University, section 9.2 and figure 9.1, has indicated that there *is* variation in the rate of use of the system; however, this data is preliminary, and studying larger samples would permit finer definition as to the exact times of student-prompted transmissions. Peaks in these transmissions can then be qualitatively explored by examining their distribution in time (perhaps even on a second by second basis). It can be quantified by examination of the correlation in time between handsets' transmissions, again using the product-moment correlation coefficient, with the null hypothesis that students' opinions are independent of each other. The statistical significance of the mean height of peaks observed in the data can be determined by comparison with the "background" level encountered at other times.

10.2.4 Exploring the nature of in-lecture learning

The handsets can also be used to study the process of student learning. If lectures are videotaped the student prompted feedback can be placed within its teaching context. This permits the exploration of learning within lecture theatres with high resolution with respect to *all* lecture events. For example, if particularly many students indicate

similar difficulties with the lecture contemporaneously it is highly probable this is due to the lecturer's actions in the preceding moments. The classification of the major features of these actions, and observation of many such peaks, would allow a falsifiable test: the observation of a peak in student-prompted handset usage and *then* the confirmation of a lecturer action. If this proved possible best practice for lecturers would become obvious, and undeniable.

10.2.5 Changing the communication possibilities

I have constrained the feedback system to a finite set of communication possibilities. Lecturers may find in practice, for example, it to be useful to have a button defined as "repeat what you just said". This possibility has been designed into the system: lecturers can use the buttons in different ways simply by defining them differently. The feedback system is a *tool* for communication.

10.2.6 Lectures as tutorials

As the richness and density of lecture communication increases it is hoped that the quality of learning achieved in lectures will approach that provided by small class teaching, such as tutorials. It may take years before enough experience with the system has been obtained to make this change. However, this possibility is the most alluring of all - the provision of the highest quality teaching to the greatest number of students without debilitating expense.

Appendix 1   Probability of collision in a random access system

The performance of a multiple transmitter, single receiver random access system is dependent on the nature of the input stream; it is assumed that it is: stationary, memoryless and orderly.[1] It will be shown that from this input stream, the distribution of transmissions obeys Poisson statistics, that is the probability of m customers arriving in a time interval of length T is:

$$P_k(T) = \frac{(\lambda T)^m}{m!} e^{-\lambda T}$$

where λ is a constant. This distribution applies to any fixed point in a transmission - for convenience, we consider the start point.[2] Finally, this distribution is applied to our system, and the relevant probability of failure, P_C, is shown to be:

$$P_C = 1 - e^{-2Nk\tau}$$

where τ is the duration of a single transmission, with N transmitters each transmitting k messages per unit time.

A1.1   Input stream

Stationary input is that the probability that m customers will arrive in the time interval (T $\Rightarrow$ T +τ) depends on τ and m only.[3] Memorylessness is that the probability of these m customers arriving in this time period, starting at time T, does not depend on the number of customers that have arrived before time T; that is, the probabilities of any given number of customers arriving in time T are mutually independent with respect to time. An input stream is orderly if it is highly unlikely that two customers will arrive at the same instant. More formally: with the probability of two or more customers arriving at the same time interval of duration *h*

---

[1] these criteria were originally demonstrated in [STEK]
[2] throughout this appendix liberal use has been made of [KLEI] and [GNEN]
[3] for any system comprising of n arbitrary length, non-overlapping time intervals of length T

being $P_{>1}(h)$ (throughout this document, a probability expressed by $P_X(t)$ is the probability that x customers will arrive during time t):

$$P_{>1}(h) \underset{\lim h \to 0}{=} 0 \qquad \text{Eq. A1.1}$$

with the additional stronger condition that as h tends towards zero the probability $P_{>1}(h)$ tends towards zero at a greater rate:

$$\frac{P_{>1}(h)}{h} \underset{\lim h \to 0}{=} 0 \qquad \text{Eq. A1.2}$$

A1.1.1 Further considerations

In the time interval $T+h$ (from an arbitrary start time T_0) there are m customer arrivals. Two further equations are used in this derivation, that for the time interval h:

$$P_0(h) = 1 - \lambda h + P_{>1}(h) \qquad \text{Eq. A1.3}$$

$$P_1(h) = \lambda h + P_{>1}(h) \qquad \text{Eq. A1.4}$$

where λ is a constant; these equations can be mathematically verified [STEK][KLEI].

A1.2 Derivation

These m arrivals can occur in (m + 1) mutually exclusive ways:

- during interval T all m customers arrive
- during interval T, m - 1 customers arrive and during interval h one customer arrives

[...]

- during interval T no customers arrive; during interval h all m customers arrive.

144

Mathematically (since memorylessness ensures there are no effects from the *order* of arrival):

$$P_m(T + h) = \sum_{j=0}^{m} P_j(T)\, P_{m-j}(h) \qquad\qquad \text{Eq. A1.5}$$

Separating this sum into three parts:

$$P_m(T + h) = P_m(T)P_0(h) + P_{m-1}(T)\,P_1(h) + \sum_{j=2}^{m} P_j(T)\, P_{m-j}(h) \qquad \text{Eq. A1.6}$$

The rightmost term is simply the probability of more than one arrival in time h and so reduces, no matter how large m becomes, to:

$$\sum_{j=2}^{m} P_j(T)\, P_{m-j}(h) = P_{>1}(h) \qquad\qquad \text{Eq. A1.7}$$

Equation A1.6 can be rewritten, using equations A1.3, A1.4 and A1.7:

$$P_m(T + h) = P_m(T)\{1 - \lambda h + P_{>1}(h)\} + P_{m-1}(T)\{\lambda h + P_{>1}(h)\} + P_{>1}(h)$$

which, subtracting $P_m(T)$ from both sides, expanding and dividing through by h becomes:

$$\frac{P_m(T + h) - P_m(T)}{h} = -\lambda P_m(T) + \frac{P_{>1}(T)}{h} P_m(T) + \lambda P_{m-1}(T) +$$

$$\frac{P_{>1}(T)}{h} P_{m-1}(T) + \frac{P_{>1}(T)}{h}$$

Now, tending h towards zero, the left hand side of the above equation becomes the formal derivative of $P_m(T)$; several terms on the right equate to zero due to equation A1.2, giving:

$$\frac{d}{dt} P_m(T) = -\lambda P_m(T) + \lambda P_{m-1}(T)$$ Eq. A1.8

This is the solution for all values of m, bar that where m is zero (since there cannot be a probability, for this stream, of *less* than 1 arrival). To approach this case we return to equation A1.5 with m = j = 0:

$$P_0(T + h) = P_0(T)P_0(h)$$

which, when combined with equation A1.4, gives directly

$$P_0(T + h) = P_0(T) (1 - \lambda h + P_{>1}(h))$$

again, subtracting $P_0(T)$ from both sides, dividing through by h and tending h towards zero gives

$$\frac{d}{dt} P_0(T) = -\lambda P_0(T)$$ Eq. A1.9

for where m = 0. Solving this is trivial, giving

$$P_0(T) = Ce^{-\lambda T}$$

Noting that the probability of no arrivals in time period of zero duration is

$$P_0(0) = 1 = Ce^{-\lambda 0}$$

$$\therefore \quad P_0(T) = e^{-\lambda T}$$ Eq. A1.10

By successive substitution of P_{m-1} into equation A1.8 one can determine that

$$P_m(T) = \frac{(\lambda T)^m}{m!} \, e^{-\lambda T}$$

Eq. A1.11

which is the equation governing Poisson statistics.

A1.3 Application to feedback system

Each individual handset's transmission, of duration τ, has a collision footprint (which would cause overlap if another transmission started during it) of 2τ:
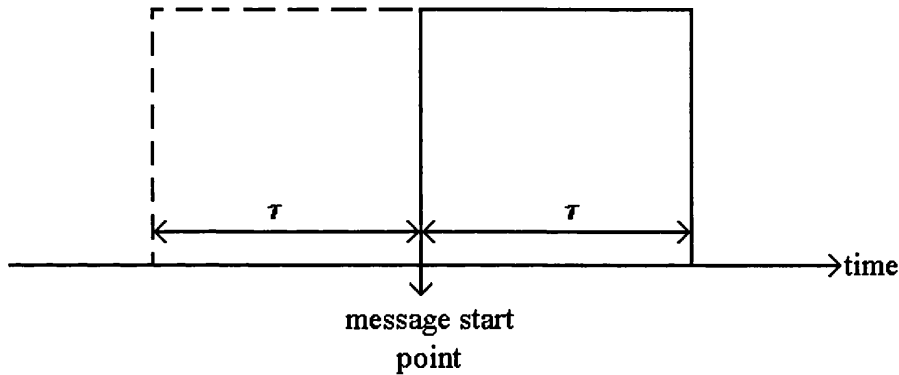
Figure A1.1 -- The collision footprint of a message of duration τ

It is assumed that the random access system input stream will be memoryless, orderly and stationary so that Poisson statistics can be applied. The probability that a message will be transmitted successfully, P_s, is the probability there will be *no* other interfering arrival in the collision footprint of a message, given that N handsets are transmitting k messages in unit time (that is, $\lambda = Nk$):

$$P_S = e^{-2Nk\tau}$$

and the corresponding probability of collision:

$$P_C = 1 - e^{-2Nk\tau}$$

This result was verified using Monte Carlo computer modelling.

147

Appendix 2   Analysis of a simple collision detection system: Aloha

Simple collision detection techniques are an initially attractive option, introducing organised network access with a requirement for energy dissipation only when a handset transmits. However, it will be demonstrated that they do not provide useful collision prevention levels at relatively low channel utilisations and hence are limited as a collision solution. This appendix examines the channel utilisation of one carrier sense system - the collision detecting Aloha system.

A2.1 System overview

The Aloha system network was a number of transmitters sending data to a single receiver. If the data was received correctly the receiver notifies the transmitter of that success. If the transmitter does not receive this notification it retransmits after a random delay until the signal from the central hub is received, or a pre-determined number of attempted retransmissions have been made.

Throughout this appendix first time attempts at sending data are labelled 'messages' and later attempts after failure(s) 'retransmissions'. There are N transmitters, sending k messages per unit time, each with a duration τ.

A2.2 Capacity utilisation

The arrival at the receiver of the messages plus retransmissions is modelled as being stationary and orderly (criteria defined in more detail in appendix 1). The input stream is also assumed to be memoryless - the probability of the arrival of messages *or* retransmissions at time T_1 is assumed to be independent of the probability at time T_2. As was demonstrated in appendix 1, this causes the arrivals to be characterised by a Poisson distribution.

Memorylessness is not an intuitively obvious assumption, as the rate of retransmission arrivals is equal to the number of previous collisions - that is, the arrival rate of data does have dependence on its history. However, "if the retransmission delay is large compared to [τ] ... and the number of retransmissions is not too large this assumption will be reasonably close to the true distribution" [ABRA]. This result was confirming by computer, using numerical methods.[1]

A2.3 Utilisation calculation

The arrival rate of starts of messages is defined to be 'r' where

$$r = k N \qquad\qquad \text{Eq. A2.1}$$

giving channel utilisation u:

$$u = r \tau \qquad\qquad \text{Eq. A2.2}$$

As can be seen, if there is perfect message alignment of messages, u equals one. If the total number of messages plus retransmissions in a stationary input stream is defined to be R, channel traffic, C, is just

$$C = R \tau \qquad\qquad \text{Eq. A2.3}$$

The probability that there will be no collision for any given message is just that of the random channel ($e^{-2R\tau}$); and the probability of collision for a message is given by equation 5.1 (p.36):

$$P_C = (1 - e^{-2R\tau}) \qquad\qquad \text{Eq. A2.4}$$

---

[1] essentially, the channel becomes unusable before the arrival rate 'memory' becomes significant

This gives the expected number of retransmissions per unit time

$$N_{RETRANSMIT,(EXPECTED)} = R(1 - e^{-2R\tau})$$ Eq. A2.5

which directly gives

$$R = r + R(1 - e^{-2R\tau})$$ Eq. A2.6

Simplifying and multiplying through by message duration

$$r\tau = R\tau e^{-2R\tau}$$ Eq. A2.7

Which, by comparison to equation A2.2 gives the relationship between channel utilisation and channel capacity:

$$u = C e^{-2C}$$ Eq. A2.8

This relationship indicates that the channel traffic will become unstable, the number of successful transmissions falling dramatically, as utilisation approaches 1/2e. This is a severe limitation: our feedback system has unconstrained data entry. If there is a period during which the demand for data transmission exceeds this limit catastrophic data loss would result. Data *would* eventually flow successfully again if there is an upper limit on the number of retransmissions the handsets can make - but the interception of an event of this nature, where a large proportion of the audience wishes to express an opinion, is precisely the design goal of the system. The use of very low channel utilisations can make this probability small; but it was felt that there were more attractive avenues for development.

Appendix 3   Software to permit data display and storage by PC

This program is written in Borland C++ programming language; comments are indicated by the notation /*comment*/. The executable program after it has been compiled can be run on any Windows PC that has Windows 3.11 or later.

```
#include <graphics.h>
#include <stdio.h>
#include <stdlib.h>
#include <conio.h>
#include <math.h>
#include <ctype.h>
#include <time.h>
#include <io.h>
#include <fcntl.h>
#include <sys\stat.h>
#include <string.h>
#define PORT_BASE (0x378)
#define PORT_STAT (PORT_BASE + 1)
#define PORT_CTRL (PORT_BASE + 2)
#define addressmask (0x78)
#define ECR_DATA (0x77A)                    /*for machine with (ECP) address with printer at 0x378)*/
#define datamask (0x07)
#define MAXNUM 5
#define MAXHANDSETS 256
#define DELAY 100
#define decay_constant 60
extern long timezone = 0;                                             /*declaring system variables*/
int gdriver = VGA;                          /*declaring variables and arrays used by the whole program*/
int gmode = VGAHI;
int i, centralvar, oldcentralvar, datapresent, value;
int countarray[MAXHANDSETS];
int graphicsarray [MAXHANDSETS], graphicstimearray [MAXHANDSETS];
double graphicsarraythree [(MAXHANDSETS * 4)];
double totalpresent;
char calendar_time[50];
char *datadir="c:\\feedback\\log\\";
char filename[255];
char fn[13];
FILE *fp;
long int start;
int run_number, time_elapsed, old_time_elapsed, updatecounter;
int portreset();                                              /*prototypes of the functions*/
int filefunctions();
int fileopening();
int timefunctions();
int reset();
int count();
int keyboard();
int threebutton();
int fivebutton();
int getdata();
int getdatadump();
int are_you_sure();
int portreturn();
int main() {
portreset();                                                  /*setting port to byte mode*/
registerbgidriver (EGAVGA_driver);                            /*initialising graphics and clearing screen*/
registerbgifont (sansserif_font);
initgraph (&gdriver, &gmode, "");
cleardevice();
filefunctions();
fileopening();
timefunctions();
reset();
getdatadump();
```

```c
count();
while (centralvar != 4){
 keyboard();
 if (oldcentralvar != centralvar){
  if (centralvar = = 2){
   count();
  }
  if (centralvar = = 1){
   centralvar = oldcentralvar;
  }
  switch (centralvar){
   case 3:
    for (i = 0; i < MAXHANDSETS; i+ +){
     graphicsarray[i] = 0;
    }
    fprintf(fp, "-3-\n");
    threebutton();
    oldcentralvar = centralvar;
    datapresent = 0;
   break;
   case 5:
    for (i = 0; i < MAXHANDSETS; i+ +){
     graphicsarray[i] = 0;
    }
    fprintf(fp, "-5-\n");
    fivebutton();
    oldcentralvar = centralvar;
    datapresent = 0;
   break;
   case 4:
   break;
  }
 }
 if (datapresent = = 1){
  if (centralvar != 4){
   switch (centralvar){
    case 3:
     threebutton();
     datapresent = 0;
    break;
    case 5:
     fivebutton();
     datapresent = 0;
    break;
   }
  }
 }
 if (centralvar != 4){                          /*to make run ok, any last bits of data are discounted*/
  getdata();                                    /*& if data is here no need to automatically update 3 button screen*/
  if(datapresent = = 1){
   updatecounter = -1;
  }
 }
 if (centralvar = = 3){                         /*every 5th loop update screen if in 3 button mode*/
  updatecounter = updatecounter+ +;            /* (i.e. decaying data) when no data is being presented*/
  if (updatecounter = = 4){
   datapresent = 1;
   updatecounter = 0;
  }
 }
 if (centralvar = = 4){
  are_you_sure();
 }
}
fclose(fp);
closegraph();
exit;
portreturn();
return(0);
}
int portreset(){                                /*setting ecp port to SPP bidirectional by 1st setting*/
value = inp( ECR_DATA );                        /*control bits 5 through 7 low, then to 001 (SPP bidirectional)*/
```

152

```c
value = ( value & 0x1F );
outp ( ECR_DATA, value | 0x70 );
return 0;
}
int filefunctions(){                                            /*creating new file to store data in*/
fp=fopen("c:\\feedback\\0lastrun.txt", "r+");          /*creating new file with filename stored in memory*/
if (fscanf(fp, "%d", &run_number) != 1){
  run_number=1;}
run_number++;
rewind(fp);
fprintf(fp, "%d", run_number);
fclose(fp);
return 0;
}
int fileopening(){                                     /*opening new file for logging data received*/
sprintf(fn, "%d.dat", run_number);
strcpy(filename, datadir);
strcat(filename, fn);
fp = fopen(filename, "w");
printf("file name %s", filename);
return 0;
}
int timefunctions(){                                   /*dealing with initial time gathering thingies*/
time_t t;
t = time(NULL);
start = t;
fprintf(fp, "The date and time of this file's start: %s\n", ctime(&t));
return 0;
}
int reset(){                                                   /*initialising entire interface*/
char mumble[80] = "Ready for the receiver to be switched";
char mumbletwo[80] = "on; please press a key when ready.";
value = inp( PORT_CTRL );                              /*putting port in reset mode*/
outp( PORT_CTRL, 0x00 );
settextstyle(SANS_SERIF_FONT, HORIZ_DIR, 0);
outtextxy(0, 173, mumble);
outtextxy(0, 208, mumbletwo);
getch();
outp(PORT_CTRL, 0x2B);                                 /*returning port to default configuration*/
cleardevice;
return 0;
}
int getdatadump(){          /*getdatadump routine, clearing the FIFO from spurious data stored on start up*/
int nomoredata;
while(!(inp(PORT_STAT) & 0x40))                        /*waiting for hardware clock*/
  ;
value = inp(PORT_CTRL);                        /*resetting hardware clock and using delay to ensure it resets*/
outp(PORT_CTRL, value | 0x01);
outp(PORT_CTRL, value & ~0x01);
for (i = DELAY; i > 0; i--)
  ;
outp(PORT_CTRL, value | 0x01);
value = inp(PORT_CTRL);                                /*acknowledge signal and delay*/
outp(PORT_CTRL, value | 0x02);
outp(PORT_CTRL, value & ~0x02);
for (i = DELAY; i > 0; i--)
  ;
nomoredata = inp(PORT_STAT) & 0x08;                    /*looking to see if nomoredata pin is high*/
while (! nomoredata){                                  /*data gathering loop*/
outp(PORT_CTRL, value | 0x02);                         /*acknowledge low*/
value = (inp(PORT_CTRL));  /*get next byte with delay to ensure receiver reacts using next-byte-please pin*/
outp(PORT_CTRL, value & ~0x04);
outp(PORT_CTRL, value | 0x04);
for (i = DELAY; i > 0; i--)
  ;
inp(PORT_BASE);                                        /*reading the data*/
outp(PORT_CTRL, value & ~0x04);                        /*ending read operation*/
value = inp(PORT_CTRL);                                /*acknowledge signal and delay*/
outp(PORT_CTRL, value | 0x02);
outp(PORT_CTRL, value & ~0x02);
for (i = DELAY; i > 0; i--)
  ;
```

```
 nomoredata = inp(PORT_STAT) & 0x08;                          /*seeing if there is more data to get*/
 }
value = inp(PORT_CTRL);                     /*resetting hardware clock and using delay to ensure it resets*/
outp(PORT_CTRL, value | 0x01);
outp(PORT_CTRL, value & ~0x01);
for (i = DELAY; i > 0; i--)
 ;
outp(PORT_CTRL, value | 0x01);
outp(PORT_CTRL, 0x2B);                                        /*set interface in default*/
return(0);
}
int count(){                                    /*Counting the number of handsets to provide 100% figure*/
int j, k, nomoredata, thedata, endloop, wan, oops;
int thisworks = 0;
char mumblethree[80] = "Please ensure those present have\n";
char mumblefour[80] = "pressed a key to register their presence.\n";
char mumblefive[80] = "When ready, please choose which display\n";
char mumblesix[80] =  "('3' or '5' button) you would like.";
char mumbleten[80] = "No-one has pressed a button yet!";
char mumbleeleven[80] = "Press a key to continue registering;";
char mumbletwenty[80] = "or press 'e' to exit program here.";
cleardevice();
settextstyle(SANS_SERIF_FONT, HORIZ_DIR, 0);
outtextxy(0, 103, mumblethree);
outtextxy(0, 138, mumblefour);
outtextxy(0, 173, mumblefive);
outtextxy(0, 208, mumblesix);
for (i = 0; i < MAXHANDSETS; i++)                            /*initialising counting array*/
 {countarray[i] = 0;}
while(!(thisworks)){                      /*hold loop to gather data until user wishes to quit registering*/
 wan = 0;
 while(!(inp(PORT_STAT) & 0x40))                             /*waiting for hardware clock*/
  ;
 value = inp(PORT_CTRL);                  /*reset hardware clock and using delay to ensure it resets*/
 outp(PORT_CTRL, value | 0x01);
 outp(PORT_CTRL, value & ~0x01);
 for (i = DELAY; i > 0; i--)
  ;
 outp(PORT_CTRL, value | 0x01);
 value = inp(PORT_CTRL);                                     /*acknowledge signal and delay*/
 outp(PORT_CTRL, value | 0x02);
 outp(PORT_CTRL, value & ~0x02);
 for (i = DELAY; i > 0; i--)
  ;
 nomoredata = inp(PORT_STAT) & 0x08;                         /*looking to see if nomoredata pin is high*/
 while (! nomoredata){                                       /*data gathering loop*/
outp(PORT_CTRL, value | 0x02);                               /*acknowledge low*/
 value = (inp(PORT_CTRL));/*get next byte with delay to ensure receiver reacts using next-byte-please pin*/
 outp(PORT_CTRL, value & ~0x04);
 outp(PORT_CTRL, value | 0x04);
 for (i = DELAY; i > 0; i--)
  ;
 thedata = inp(PORT_BASE);                                   /*reading the data*/
 totalpresent = 1;                          /*acknowleding that some data has been entered*/
 j = ((thedata & addressmask) >> 3);          /*putting the address into the register of those present*/
 if (countarray[j] != 1)
  {countarray[j] = 1;}
 outp(PORT_CTRL, value & ~0x04);                             /*ending read operation*/
 value = inp(PORT_CTRL);                                     /*acknowledge signal and delay*/
 outp(PORT_CTRL, value | 0x02);
 outp(PORT_CTRL, value & ~0x02);
 for (i = DELAY; i > 0; i--)
  ;
 nomoredata = inp(PORT_STAT) & 0x08;                         /*seeing if there is more data to get*/
 }
value = inp(PORT_CTRL);                     /*reset hardware clock and using delay to ensure it resets*/
outp(PORT_CTRL, value | 0x01);
outp(PORT_CTRL, value & ~0x01);
for (i = DELAY; i > 0; i--)
 ;
outp(PORT_CTRL, value | 0x01);
outp(PORT_CTRL, 0x2B);                                       /*set interface in default*/
```

```c
/* This section detects if there's been a keyboard input, and ignores all but the last keystroke. If this is    */
/* valid (and there are actually handsets present) this exits the datagathering loop.                           */
if (!kbhit())
 {wan = 1;}
while(!wan){
 printf("Invalid keyboard input\n");                                           /*test code*/
 centralvar = getch();
 if (!kbhit())
  {wan = 1;}
}
 switch (centralvar){                                         /*putting standard value onto centralvar*/
  case 51: centralvar = 3;
           thisworks = 1;
           break;
  case 53: centralvar = 5;
           thisworks = 1;
           break;
  default: centralvar = 0;
           thisworks = 0;
           break;
 }
 if ((totalpresent < 1) && (thisworks == 1)){    /*if no-one pressed button, prevent downstream problems*/
  cleardevice();
  outtextxy(0, 173, mumbleten);
  outtextxy(0, 208, mumbleeleven);
  outtextxy(0, 243, mumbletwenty);
  while(! kbhit());
  oops = getch();
  switch(oops){
   case 101: case 69:
   totalpresent = 1;
   thisworks = 1;
   centralvar = 4;
   oops = 1000;
   break;
   default:
   break;
  }
  if (oops != 1000){
   cleardevice();
   outtextxy(0, 103, mumblethree);
   outtextxy(0, 138, mumblefour);
   outtextxy(0, 173, mumblefive);
   outtextxy(0, 208, mumblesix);
   thisworks = 0;
  }
 }
}
 totalpresent = 0.0;    /*count number of addresses that have been active since beginning of this function*/
 for (i = 0; i < MAXHANDSETS; i++)
  {if (countarray[i] > 0){
    totalpresent = (totalpresent + 1);
    fprintf(fp, "%d", i);
  }
kbhit();
getch;
 }
fprintf(fp, "\n");                                          /*some break lines between this and the data*/
fprintf(fp, "\n Above handsets present.   ");
fprintf(fp, "File format: time noted, then address then data\n");
return 0;
}
int keyboard(){                                          /*Seeing if there is keyboard input at all*/
int wan = 0;
int twerp = 0;
char mumbleeight[80] = "Please enter valid input.";
char mumblenine[80] = "The program will stop until you do!";
char mumbleten[80] = "('r'eset; 'e'nd; '3' or '5' buttons)";
settextstyle(SANS_SERIF_FONT, HORIZ_DIR, 0);
if (!kbhit())
 {wan = 1;}
if (!wan){
```

```c
  centralvar = 0;
  while (!((centralvar > = 1) && (centralvar < = 5))){
   while (!wan){
    centralvar = getch();
    if (!kbhit())
    {wan = 1;}
   }
  twerp = twerp + +;
  wan = 0;
  switch (centralvar){
    case 51:
     centralvar = 3;
     break;
    case 53:
     centralvar = 5;
     break;
    case 82: case 114:
     centralvar = 2;
     break;
    case 69: case 101:
     centralvar = 4;
     break;
    case 90: case 122:
     centralvar = 1;
     break;
    default:
     centralvar = 0;
     switch (twerp){
      case 1:
      outtextxy(0, 183, mumbleeight);
      datapresent = 1;
      break;
      case 2:
      outtextxy(0, 223, mumblenine);
      datapresent = 1;
      break;
      default:
      outtextxy(0, 263, mumbleten);
      datapresent = 1;
      break;
      }
   }
  }
 }
 return (0);
 }
 int threebutton(){                              /*three button display routine WHERE OPINIONS DECAY*/
 int numbars = 3;
 char maintitle[80], subtitle[80], ytitle[80], buffer[40];
 char errmssge[80] = "Please restart program.";
 char spec_error[80];
 char *members[] = {"too slow", "about right", "too fast"};
 char nodata[80] = "(no data at the moment)";
 double values[MAXNUM], pcvalues[MAXNUM];
 int maxvalue, xspace, xmarker, i, z, blue;
 div_t a;
 double y;
 double totalvalue;
 double ageof;
 maxvalue = 100;
 time_t t;                                       /*getting current time for decay purposes*/
 t = time(NULL);
 time_elapsed = (t - start);
 cleardevice();
 rectangle (0,0,639,479);
 sprintf (maintitle, "3 buttons");
 sprintf (ytitle, "Percent");
 for (i = 0; i < MAXNUM; i + +){                          /* zeroing arrays for display */
  values[i] = 0.0;}
 for (i = 0; i < MAXHANDSETS; i + +)             /* counting data entered by only those handsets present*/
 {
  blue = graphicsarray[i];                               /* blind variable for this loop*/
```

```c
  if ((blue > 0) && (blue < 4)){
   ageof = time_elapsed - graphicstimearray[i];
   if ((ageof <=70) && (ageof >= 0)){
    graphicsarraythree[(i + blue)] = (((-1 * ageof) / decay_constant) + 1);
    if (graphicsarraythree[(i + blue)] < 0){
     graphicsarraythree[(i + blue)] = 0;
    }
    values[(blue - 1)] = values[(blue - 1)] + graphicsarraythree[i + blue];
   }
  }
 }
for (i = 0; i < numbars; i++)                              /*error protection*/
 {if (values[i] < 0)
   {settextstyle(DEFAULT_FONT, HORIZ_DIR, 2);
    sprintf (spec_error, "Abnormal data error,");
    outtextxy(163, 100, spec_error);
    outtextxy(163, 115, errmssge);
    values[i] = (-1 * values[i]);}
   }
totalvalue = 0;                    /*to let the lecturer know the program is alive if there is no data present*/
for (i = 0; i < numbars; i++)
{totalvalue = (totalvalue + values[i]);}
if (totalvalue == 0)
 {settextstyle (DEFAULT_FONT, HORIZ_DIR, 0);
  settextjustify (CENTER_TEXT, CENTER_TEXT);
  outtextxy (320, 60, nodata);
  totalvalue = 1;}
for (i = 0; i < numbars; i++)
{
pcvalues[i] = ((values[i] / totalpresent) * 100);
}
settextstyle (DEFAULT_FONT, HORIZ_DIR, 2);
settextjustify (CENTER_TEXT, CENTER_TEXT);
outtextxy (320, 40, maintitle);
settextstyle (DEFAULT_FONT, VERT_DIR, 1);
outtextxy (10, 240, ytitle);
rectangle (60, 90, 620, 440);
line (60, 354, 620, 354);
line (60, 265, 620, 265);
line (60, 178, 620, 178);
settextstyle (DEFAULT_FONT, HORIZ_DIR, 1);
settextjustify (RIGHT_TEXT, CENTER_TEXT);
sprintf (buffer, "%d", maxvalue);
outtextxy (55, 90, buffer);
a = div (maxvalue * 3,4);
sprintf (buffer, "%d", a.quot);
outtextxy (55, 178, buffer);
a = div (maxvalue , 2);
sprintf (buffer, "%d", a.quot);
outtextxy (55, 265, buffer);
a = div (maxvalue, 4);
sprintf (buffer, "%d", a.quot);
outtextxy (55, 354, buffer);
a = div (560, numbars + 1);
xspace = a.quot;
xmarker = 60;
settextjustify (CENTER_TEXT, TOP_TEXT);
for (i = 0; i < numbars; i++)
{
switch (i) {
 case 0: z = 2;
 break;
 case 1: z = 7;
 break;
 case 2: z = 4;
 break;
 }
setfillstyle (SOLID_FILL, z);
xmarker = xmarker + xspace;
outtextxy (xmarker, 445, members[i]);
y = 3.5 * pcvalues[i];
bar (xmarker - 27, 440 - y, xmarker + 27, 440);
```

157

```c
        }
        settextjustify(LEFT_TEXT, TOP_TEXT);
        return 0;
        }
        int fivebutton(){                                              /*five button display routine*/
        int numbars = 5;
        double totalvalue = 0;
        char maintitle[80], subtitle[80], xtitle[80], ytitle[80], buffer[40];
        char *members[] = {"A", "B", "C", "D", "E"};
        char errmssge[80] = "Please restart program.";
        char nodata[80] = "(no data at the moment)";
        char spec_error[80];
        double values[MAXNUM], pcvalues[5];
        int maxvalue, xspace, xmarker, z;
        div_t a;
        double y;
        maxvalue = 100;
        cleardevice();
        rectangle (0,0,639,479);
        sprintf (maintitle, "5 buttons");
        sprintf (xtitle, "Button");
        sprintf (ytitle, "Percent");
        for (i = 0; i < MAXNUM; i++){                                  /* zeroing arrays for display */
         values[i] = 0;}
        for (i = 0; i < MAXHANDSETS; i++){      /* counting data entered by only those handsets present*/
         if (countarray[i] != 0){
          switch (graphicsarray[i]){                                   /*now reading in the data for display*/
          case 1:
          values[0] = values[0]++;
          break;
          case 2:
          values[1] = values[1]++;
          break;
          case 3:
          values[2] = values[2]++;
          break;
          case 4:
          values[3] = values[3]++;
          break;
          case 5:
          values[4] = values[4]++;
          break;
         }
        }
        }
        for (i = 0; i < numbars; i++)                                  /*error protection*/
        {if (values[i] < 0)
          {settextstyle(DEFAULT_FONT, HORIZ_DIR, 2);
           sprintf (spec_error, "Abnormal data error,");
           outtextxy(163, 100, spec_error);
           outtextxy(163, 115, errmssge);
           values[i] = (-1 * values[i]);}
         }
        totalvalue = 0;                                               /*to let the lecturer know the program is alive*/
        for (i = 0; i < numbars; i++)
        {totalvalue = (totalvalue + values[i]);}
        if (totalvalue == 0)
         {settextstyle (DEFAULT_FONT, HORIZ_DIR, 0);
          settextjustify (CENTER_TEXT, CENTER_TEXT);
          outtextxy (320, 60, nodata);
          totalvalue = 1;}
        for (i = 0; i < numbars; i++)
        {
        pcvalues[i] = ((values[i] / totalpresent) * 100);
        }
        settextstyle (DEFAULT_FONT, HORIZ_DIR, 2);
        settextjustify (CENTER_TEXT, CENTER_TEXT);
        outtextxy (320, 40, maintitle);
        settextstyle (DEFAULT_FONT, HORIZ_DIR, 1);
        outtextxy (320, 460, xtitle);
        settextstyle (DEFAULT_FONT, VERT_DIR, 1);
        outtextxy (10, 240, ytitle);
```

158

```
rectangle (60, 90, 620, 440);
line (60, 354, 620, 354);
line (60, 265, 620, 265);
line (60, 178, 620, 178);
settextstyle (DEFAULT_FONT, HORIZ_DIR, 1);
settextjustify (RIGHT_TEXT, CENTER_TEXT);
sprintf (buffer, "%d", maxvalue);
outtextxy (55, 90, buffer);
a = div (maxvalue * 3,4);
sprintf (buffer, "%d", a.quot);
outtextxy (55, 178, buffer);
a = div (maxvalue , 2);
sprintf (buffer, "%d", a.quot);
outtextxy (55, 265, buffer);
a = div (maxvalue, 4);
sprintf (buffer, "%d", a.quot);
outtextxy (55, 354, buffer);
a = div (560, numbars + 1);
xspace = a.quot;
xmarker = 60;
settextjustify (CENTER_TEXT, TOP_TEXT);
for (i = 0; i < numbars; i++)
{
 switch (i) {
  case 0: z = 9;
  break;
  case 1: z = 6;
  break;
  case 2: z = 14;
  break;
  case 3: z = 5;
  break;
  case 4: z = 3;
  break;
    }
 setfillstyle (SOLID_FILL, z);
 xmarker = xmarker + xspace;
 outtextxy (xmarker, 445, members[i]);
 y = 3.5 * pcvalues[i];
 bar (xmarker - 27, 440 - y, xmarker + 27, 440);
}
settextjustify(LEFT_TEXT,TOP_TEXT);
return 0;
}
int getdata(){                                                  /*data gathering routine*/
int j, k, nomoredata, thedata;
while(!(inp(PORT_STAT) & 0x40))                                 /*waiting for hardware clock*/
 ;
value = inp(PORT_CTRL);                     /*resetting hardware clock and using delay to ensure it resets*/
outp(PORT_CTRL, value | 0x01);
outp(PORT_CTRL, value & ~0x01);
for (i = DELAY; i > 0; i--)
 ;
outp(PORT_CTRL, value | 0x01);
value = inp(PORT_CTRL);                                         /*acknowledge signal and delay*/
outp(PORT_CTRL, value | 0x02);
outp(PORT_CTRL, value & ~0x02);
for (i = DELAY; i > 0; i--)
 ;
nomoredata = inp(PORT_STAT) & 0x08;                    /*looking to see if nomoredata pin is high*/
if (!nomoredata){                                      /*if there is data, getting the time*/
 time_t t;
 t = time(NULL);
 time_elapsed = (t - start);
}
if ((! nomoredata) && (old_time_elapsed != time_elapsed)){    /*if data present, record time it was gathered*/
 fprintf(fp, "@ %d seconds: ", time_elapsed);
 old_time_elapsed = time_elapsed;
}
while (! nomoredata){                                           /*data gathering loop*/
 datapresent = 1;                          /*setting datapresent to 1 to indicate this fact to rest of program*/
 outp(PORT_CTRL, value | 0x02);                                /*acknowledge low*/
```

159

```c
/* getting next byte with delay to ensure receiver reacts using next-byte-please pin;        */
/* abandoned use of next-byte-here pin                                                        */
value = (inp(PORT_CTRL));
outp(PORT_CTRL, value & ~0x04);
outp(PORT_CTRL, value | 0x04);
for (i = DELAY; i > 0; i--)
  ;
thedata = inp(PORT_BASE);                                                /*reading the data*/
j = ((thedata & addressmask) >> 3);    /*putting the data into the graphics array, and the time of input*/
k = (thedata & datamask);
graphicsarray[j] = k;
graphicstimearray[j] = time_elapsed;


/*        Remmed out for the moment - can use a temporary memory store                */
/*        - if a machine has smartdrv this is not used.  As smartdrv is present        */
/*        on the machine that this program has been compiled for this is not currently used.  */
/*        Left in for completeness.                                                    */
/*memstorearray[memstore] = thedata;                                                   */
/*memstore = memstore + +;                                                             */

fprintf(fp, "%8d %5d\n", j, k);            /*directly storing this in the file, smartdrv will take care of it*/
outp(PORT_CTRL, value & ~0x04);                                      /*ending read operation*/
value = inp(PORT_CTRL);                                      /*acknowledge signal and delay*/
outp(PORT_CTRL, value | 0x02);
outp(PORT_CTRL, value & ~0x02);
for (i = DELAY; i > 0; i--)
  ;
nomoredata = inp(PORT_STAT) & 0x08;                          /*seeing if there is more data to get*/
  }
value = inp(PORT_CTRL);                     /*resetting hardware clock and using delay to ensure it resets*/
outp(PORT_CTRL, value | 0x01);
outp(PORT_CTRL, value & ~0x01);
for (i = DELAY; i > 0; i--)
  ;
outp(PORT_CTRL, value | 0x01);
outp(PORT_CTRL, 0x2B);                                               /*set interface in default*/
return(0);
}
int are_you_sure(){                              /*"checking whether user wishes to exit" function*/
int whathit;
char mumbletwelve[80] =   "You have pressed the exit button. Is\n";
char mumblethirteen[80] = "this correct? Press 'y' to confirm or\n";
char mumblefourteen[80] = "any other key to return to the program.";
cleardevice();
settextstyle(SANS_SERIF_FONT, HORIZ_DIR, 0);
outtextxy(0, 183, mumbletwelve);
outtextxy(0, 223, mumblethirteen);
outtextxy(0, 263, mumblefourteen);
whathit = getch();
switch (whathit){
 case 89: case 121:
 break;
 default:
 centralvar = oldcentralvar;
 datapresent = 1;
 break;
}
return 0;
}
int portreturn(){
value = inp( ECR_DATA );                                             /*returns port to ECP mode*/
value = ( value & 0x1F );
outp ( ECR_DATA, value | 0xC0);
return 0;
}
```

160

Appendix 4   System 2 radio transmitter and receiver

A4.1 Radio transmitter

The radio transmitter used was an RF9901 [RF99t] frequency shift keying IC,
operating at 916.36 MHz, figure A4.1.
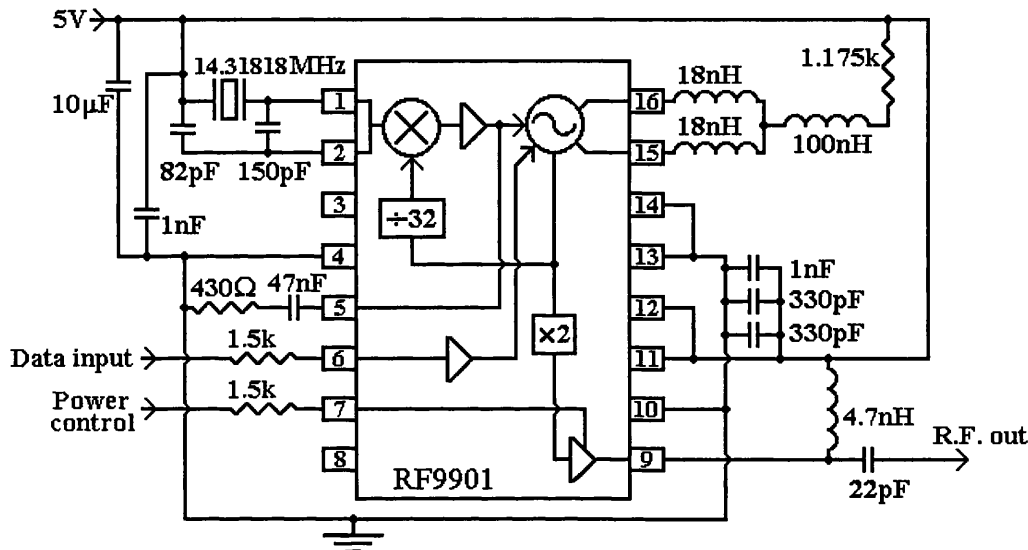
*Figure A4.1 -- The circuit diagram of the system 2 radio transmitter, including an
indicative schematic of the RF9901 integrated circuit*

A4.1.1 Pin description

The function of each pin is described in table A4.1.

Pin number	Function
1 & 2	14.31818MHz input from crystal oscillator
3 & 8	Not used in this circuit
4, 10, 13 & 14	Ground
5	Phase lock loop filter
6	Data input
7	Power control
9	916MHz output
11 & 12	5 Volt supply
15 & 16	Voltage Controlled Oscillator (VCO) resonator power inputs

Table A4.1 -- The pin descriptions of the RF9901

A4.1.2 Functional description

A stable crystal oscillator controlled 14.31818MHz input signal is produced by pins 1 and 2; this signal is passed through a phase comparator amplified and applied to a voltage controlled oscillator (VCO). This VCO, powered via pins 15 and 16, begins to resonate; this signal is divided by 32 and compared in phase with the crystal oscillator input frequency. The resultant voltage from the *phase* difference between these two frequencies is again passed through the amplifier. High frequency phase differences result in a high frequency voltage changes from the phase comparator, and these voltage components are passed straight to ground via the loop filter, pin 5. The lower frequency phase differences from the phase comparator charge the capacitor, and so are applied to the VCO, changing its output frequency. Therefore, over time, the VCO is phase-locked to a frequency 32 times that of the crystal oscillator, 458.18MHz.

The modulation input from pin 6 is amplified and applied to VCO and causes a change in the output frequency of the VCO. This frequency change is also divided by 32, compared in phase to the crystal oscillator signal and amplified. The resultant phase comparator voltage will *not* be applied to the VCO if the modulation has a high enough frequency so that the capacitor on pin 5 does not charge. However, if the modulation input voltage remains constant for too long it will cause a phase difference output from the phase comparator of sufficient duration to charge the capacitor sufficiently to be applied to the VCO - and the modulation will begin to be "corrected". Therefore there is a maximum duration of any stable modulation input.

During power up, whilst the VCO initially stabilises at 458MHz over some milliseconds, no modulation voltage should be applied to pin 6. After this period, modulation can begin - but the low pass filter must be adjusted such that the maximum expected stable modulation duration is not "corrected". This implies a decrease in the bandwidth of the filter - a decrease in the upper limit of those frequencies applied to the VCO via the maximisation of the low pass filter

capacitance. However, during initial (power-on) VCO stabilisation these higher frequencies will exhibit less aliasing of the voltage resulting from the phase comparison - their 'sampling' frequency will be closer to the true phase difference between the (divided by 32) VCO output and the crystal oscillator input. These higher frequencies will speed the process of loop locking. Therefore, it is desirable to have the bandwidth of the low pass filter maximised, by increasing the upper limit to those frequencies applied to the VCO via *minimisation* of the low pass filter capacitance. The values of the capacitor and resistor were experimentally verified to be the best compromise for our system between these conflicting demands.

The VCO output (plus frequency deviation due to modulation) is then frequency doubled to the correct output frequency (916.36MHz). Finally, it is passed through a power amplifier. This amplifier can be disabled by the application of zero volts to pin 7 (the output level is a maximum -10dBm with the amplifier disabled, and +4dBm when it is on). During VCO stabilisation the power control input is held low to minimise both transmission duration and power dissipation.[1] The 916MHz signal is then output from pin 9 to a quarter-wavelength monopole aerial.

A4.2 Radio receiver

The receiver used was the matched RF9902 [RF99R], figure A4.2.

---

[1] initial VCO stabilisation requires several milliseconds; the RF9901 dissipates 22mA with the output amplifier disabled, and 28mA when it is active
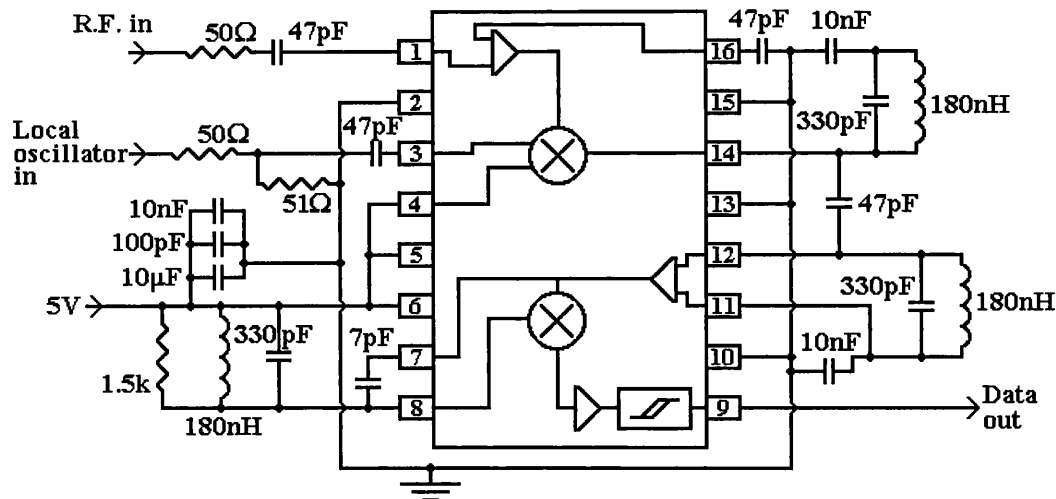
Figure A4.2 -- The circuit diagram of the system 2 radio receiver, including an indicative schematic of the RF9902 integrated circuit

A4.2.1 Pin description

The function of each pin is described in table A4.2.

Pin number	Function
1	Radio frequency input
2, 10, 13 & 15	Ground
3	Local oscillator input
4, 5 & 6	5 Volt supply
7	Intermediate signal output to quadrature tank
8	Quadrature tank (resonant at intermediate frequency)
9	Baseband data out (at CMOS voltage levels)
11 & 12	Intermediate frequency (IF) inputs (pin 11 the complementary input)
14	IF mixer output
16	Radio frequency ground (out-of-phase input)

Table A4.2 -- The pin descriptions of the RF9902

A4.2.2 Functional description

The radio frequency signal from the quarter wavelength monopole aerial is input to the RF9902 and amplified. It is then mixed with a local oscillator signal. This local oscillator is a RF9901 transmitter, configured as figure A4.1 but with a 14.000MHz

crystal oscillator,[2] providing an 896MHz sine-wave input. The output of the mixer is a 20.36MHz intermediate frequency (IF) signal, plus the frequency deviations due to data. This input is passed, via an AC connection, to an IF amplifier. The signal is then split, with one component being output to a resonant circuit and the other component being passed to a mixer. This 'quadrature tank circuit' resonates at the IF, with a 90° phase shift between its resonance and the central IF frequency. The output from the quadrature tank circuit is mixed with the IF, removing the carrier component and, using the relative phase shift versus frequency, the modulation state of the signal is output. The baseband data stream is then amplified, and passed through a Schmitt Trigger which exhibits hysteresis and so provides 'clean' transitions between modulation states and outputs at CMOS voltage levels.

---

[2] without any modulation input and with the output amplifier permanently on

Appendix 5   System 2 receiver

The following diagrams detail the system 2 receiver. Figure A5.1 - which is the initial and tight checks, with clock synchronisation - includes some standard circuit elements. In the 'initial check' circuitry, the rising and falling edge detectors are effected using the circuit diagrams of [HORO p. 557, circuits A and B]. The start/stop interval and interval counter functions are constructed according to [HORO p.1023, figure 15.31]. The logic for the initial check circuitry simply permits the latch (a simple SR latch) to be clocked if the counter has counted 31, 32 or 33 clock pulses. The -5V power supply is that of [MAX7b], application circuit. The integrator is effected as per [HORO, p.222, figure 4.48] with a 10k leading resistor. The 'reset' is provided by a DG419 analogue switch connected between the output and input [DG41]. The integrator's output can be fed into any standard comparator with the 5 volt power supply as a reference voltage.
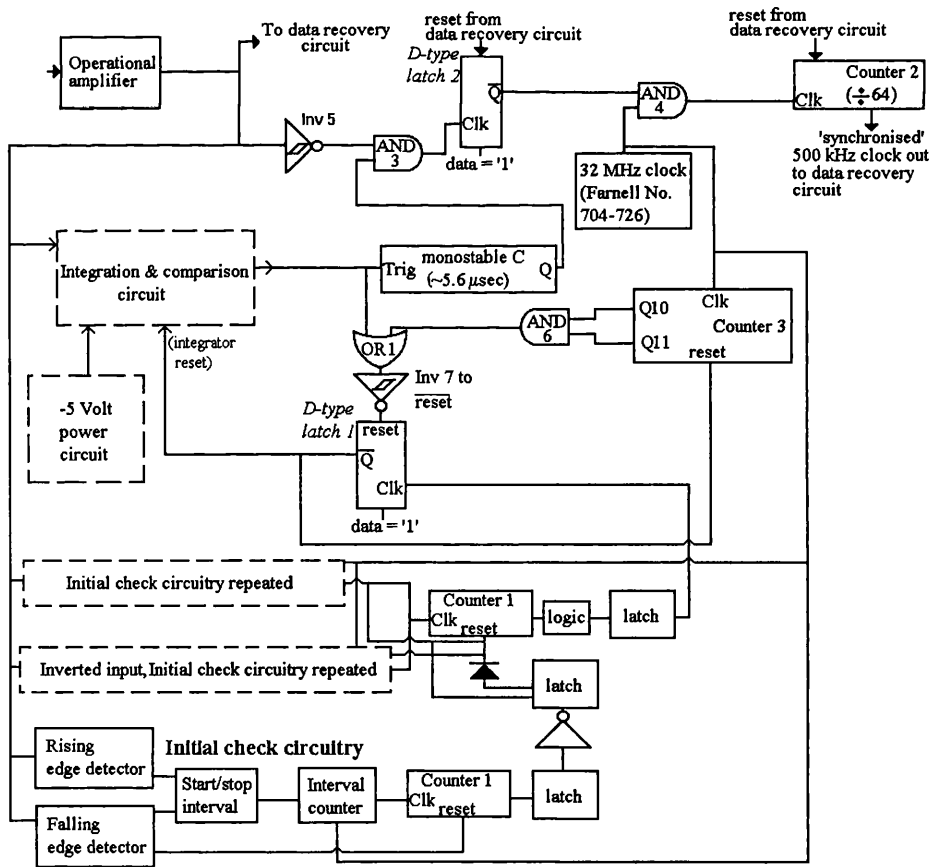

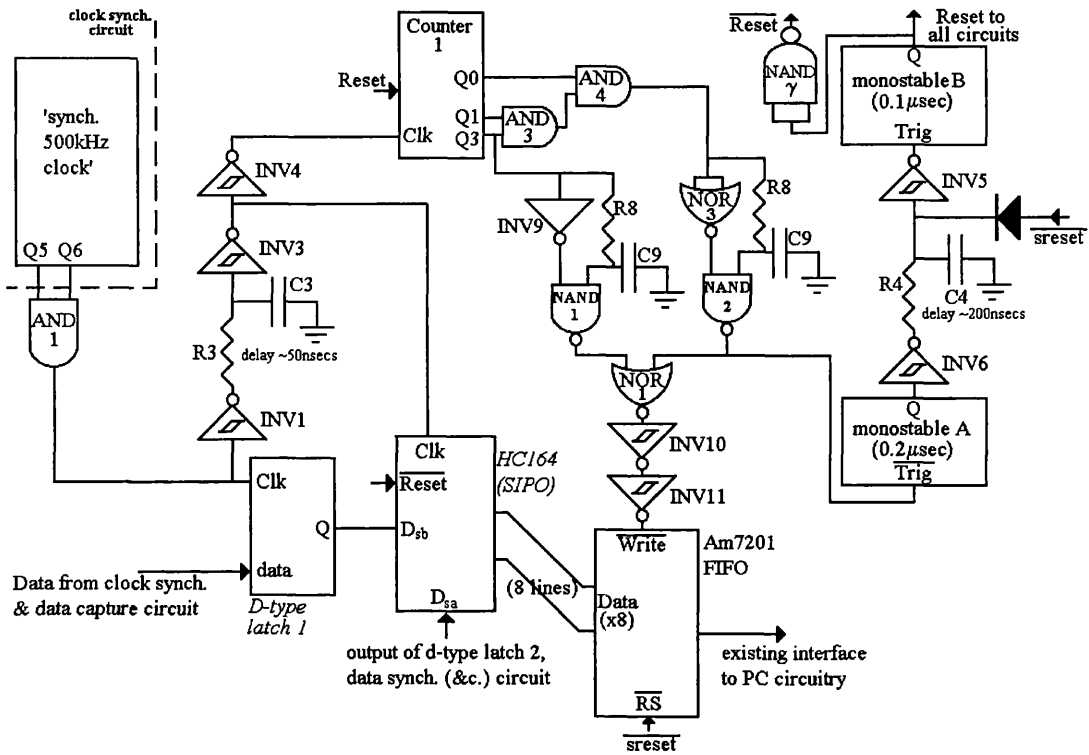*Figure A5.1 -- Initial check, tight check and clock synchronisation*

Figure A5.2 -- Data capture to FIFO storage

167

Appendix 6   Handset registration

When the system is activated, and at other times as determined by the lecturer, it is necessary to enumerate the handsets present to permit this total to be set as 100% of the audience display: 'registration'. If a collision, or other data loss through the effects of noise, occurs a handset may not be registered resulting in an inaccurate display to the lecturer. With a few (less than ten) handsets this non-registration is simple to discern and correct. When audience size rises much above this it is more difficult to identify and correct any problems. The simplest means by which to ensure that all handsets present are registered is to require them to repeatedly transmit the receipt of one transmission which will cause its address to be registered. However, when the number of handsets in use is less than the total number of handsets provided by the system there is a probability that a handset will suffer a bit error in its address causing it to register an address that is not present. It is of interest to ensure that the number of times a handset transmits during registration is such that it is extremely likely that its address will be registered whilst minimising the probability that a bit error will cause a false address to be registered (a 'ghost' registration). The following calculation reveals that optimal registration requires each handset to transmit three times, which virtually assures its the registration of its address whilst retaining a probability of 85% that there will be one or zero ghost registrations.

A6.1 Definitions

The handset system comprises of N handsets, of which x may be used in a lecture. Each handset is required to transmit y times during registration. The probability of data loss through collision is α; the probability of bit errors causing an address to be erroneously received is β; the probability of transmission loss through the effects of noise is γ.

Any handset transmission might be affected by some or all of the above probabilities - the possible outcomes are listed in table A6.1.

Option number	Transmission lost due to collision (α)	Address received erroneously (β)	Transmission lost due to noise (γ)
1	Y	Y	Y
2	Y	Y	N
3	Y	N	Y
4	Y	N	N
5	N	Y	Y
6	N	Y	N
7	N	N	Y
8	N	N	N

Table A6.1 -- The possible outcomes of a handset transmission

A6.2 Probability of all handsets being registered

The only option of table A6.1 resulting in the correct registration of a handset transmission is option 8. Designating the probability of a correct registration of a single handset in an x handset system, where these handsets transmit once P(r,1,1):

$$P(r,1,1) = (1 - \alpha)(1-\beta)(1-\gamma) = \rho \qquad \text{Eq. A6.1}$$

Therefore the probability of this handset *not* registering is

$$P(\text{not-}r,1,1) = (1 - \rho) \qquad \text{Eq. A6.2}$$

For that one handset transmitting y times the only chance of not registering is if *all* transmissions failed

$$P(\text{not-}r,1,y) = (1 - \rho)^y \qquad \text{Eq. A6.3}$$

So the probability of registering is

$$P(r,1,y) = \{1 - (1 - \rho)^y\} \qquad \text{Eq. A6.4}$$

The probability of all x handsets registering after y transmissions per handset is simply the probability that all succeeded

$$P(r,x,y) = \{1 - (1 - \rho)^y\}^x \qquad\qquad \text{Eq. A6.5}$$

A6.3 Probability of no ghost registrations

The probability of a false registration by one handset in one trial is designated $P(f,1,1)$; this can only come about if option 6 of table A6.1 occurs.

$$P(f,1,1) = (1-\alpha)\beta(1-\gamma) = \phi \qquad\qquad \text{Eq. A6.6}$$

However, it is of no concern if this bit probability causes the handset's wrong address to be one that is present in the lecture - I only care if the created handset address is in the storage box. This probability is designated $P(F,1,1)$. The total number of different addresses that a bit error could cause is $(N - 1)$. The total number of handsets not present is $(N - x)$. Therefore the probability of a ghost registration is

$$P(F,1,1) = \frac{(N-x)}{(N-1)} \phi = D\,\phi \qquad\qquad \text{Eq. A6.7}$$

The probability that a handset will not give a ghost address is:

$$P(\text{not-}F,1,1) = (1 - D\phi) \qquad\qquad \text{Eq. A6.8}$$

The probability that a handset will not give a ghost registration in y trials is the probability it does not in any of them

$$P(\text{not-}F,1,y) = (1 - D\phi)^y \qquad\qquad \text{Eq. A6.9}$$

The probability that no ghost registrations occur in a system of x handsets is

$$P(\text{not-F},x,y) = ((1 - D\phi)^y)^x = (1 - D\phi)^{yx} \qquad\qquad \text{Eq. A6.10}$$

A6.4 Probability of a single ghost registration

If only a single ghost registration occurs it will be noticeable if there are only a few (less than ten) handsets in use - and will not cause a significant display anomaly if many are present (a few percent difference in the height of the display of each student's opinion). The probability of a single ghost registration from a single handset transmission is given by equation A6.7; the probability of only 1 ghost from y transmissions from this handset is

$$P(\text{one-F},1,y) = (^{y}C_1)\,(D\phi)\,(1-D\phi)^{(y-1)} \qquad\qquad \text{Eq. A6.11}$$

The probability that there will be one ghost registration from a system of x handsets is the probability of equation A6.11 multiplied by the probability that all of the others have *not* had a ghost registration

$$P(\text{one-F},x,y) = (^{x}C_1)\;\{(^{y}C_1)\,(D\phi)\,(1-D\phi)^{(y-1)}\}\;\{1-D\phi\}^{(x-1)y}$$
$$= xy\,(D\phi)\,(1-D\phi)^{(x-1)y + (y-1)}$$
$$= xy\,(D\phi)\,(1-D\phi)^{xy-1} \qquad\qquad \text{Eq. A6.12}$$

A6.5 Results

The probability of a bit error prior to clock recovery, α, or a bit error in the address, β, is given by equation 8.1 (p.112) - the number of bits prior to clock recovery is a maximum 26, the number of bits in the address is 8. The bit error probability is assumed to be the maximum 5×10^{-4}. The probability of transmission loss due to collision, γ, is given by equation 5.1. The highest probability of collision occurs when all handsets transmit within 1 second, 2.0%; this probability can fall close to zero as it is unlikely that the students will arrive at the start of a lecture and register within even one minute.

Equations A6.5, A6.10 and A6.12 confirm the intuitive assumption that it would be best to have the handsets transmit their address enough, but no more, times than necessary to virtually ensure their registration. Computer modelling of these equations indicates that under the conditions where transmission loss is worst - when all the system's handsets transmit in one second - three button presses suffice with the probability of one or zero ghost registrations a maximum of 85%. Therefore 3 button presses will be required for the handset registration process.

Appendix 7   System 2 handset cost

The following costings are from [FARN] apart from where marked in column B. In each case the cost for each component is included at two levels: the cost where a single system is being constructed, and the lower cost where several systems are being built and economies of scale apply.

Table A7.1 is a breakdown of the costs of obtaining components for each system 2 handset. Column A is the functional description of the component; column B its part number, in italics, the number of components per part where more than one is provided; column C is the total number of components used in the handset with the number of parts in italics where different. Column D is the total cost of the parts (the number of parts needed multiplied by the cost per part) if only a single system is built, in pounds sterling; column E the total cost of parts if building multiple systems (without the use of an ASIC and in volumes numbering less than one thousand handsets in total).

The total cost for of components in constructing a single handset where a single system is being constructed is £50.94. Total cost of components in constructing a single handset where multiple systems are being built is £38.93. These figures presume that the 3V power supply has not been used for economy, as the two 1.5V 'AA' batteries will supply this voltage without DC-DC conversion (the current handset does not use a photovoltaic power source).

To this should be added the costs of construction itself. The cost of producing a mask for a printed circuit board are £35; this capital cost can be neglected if over a thousand boards are made. The cost of the production of each board is £1.00 [RAED]. The cost of fixing the components to each board are £1.00 per board [ibid.].

The total cost per handset when multiple systems are constructed is **£40.93**

A	B	C	D	E
HC4040	MM74HC4040N	2	0.46	0.36
D-type latch	MM74HC74AN *(2)*	5 *(3)*	0.42	0.30
NOR gate	MM74HC02N *(4)*	15 *(4)*	0.44	0.36
Shift register	MM74HC165N	5	0.85	0.80
AND gates	MM74HC08N *(4)*	3 *(1)*	0.10	0.09
Triple input NOR gates	MM74HC27N *(3)*	3 *(1)*	0.21	0.16
DG419	DG419	1	1.78	1.09
555 timer	ICM7555IPA	1	0.32	0.24
Resistors	SFR16S series	34	1.29	0.47
Capacitors (10μF)	TAP series	2	0.30	0.26
Capacitors (<0.5μF)	Ceramic CW series	26	3.12	3.12
Inductor (22μH) (L1 on board 1)	B82111-E-C22	1	0.59	0.39
Surface mount (SMT) resistors (transmitter board)	P-FCT series [DIGI]	4	0.40	0.25
SMT capacitors (transmitter board)	PCC-GCT series [DIGI]	9	0.90	0.60
SMT inductors (transmitter board)	0805 series [COIL]	4	3.20	2.40
14.318 MHZ crystal	SE3429-ND	1	0.64	0.55
1 MHz crystal	AEL1MHz	1	1.51	1.30
PVD1354 photovoltaic relay	PVD1354	2	6.52	5.72
Junction transistors	2N5089 discontinued. Replacement is BC549C	6	0.14	0.11
FETs	2N7000(SLX)	7	0.49	0.42
Single Pole Double Throw switches	SP35P)-0-02-76-71 [APEM]	5	8.75	4.75
MAX756 (5V output DC-DC converter)	MAX756CPA	1	2.63	2.17
HA7210 oscillator module	HA7210	1	0.14	0.10
Power diode (for MAX756) (D2)	1N5817	1	0.46	0.17
Signal diode (D1)	BAT43	1	0.08	0.05
RF9901 transmitter	RF9901	1	10.18	8.70
Enclosure	SQ2-ModAB1.5 [TEKO]	1	5.02	4.00

Table A7.1 -- The component costs of a single handset

174

Appendix 8   Energy dissipated by 1 system 2 handset transmission

This appendix first considers the power dissipation from the 3V supplied components, and then the 5V supplied components. All calculations are conducted in SI units, with the resultant energy consumption quoted as μJ.

A8.1 Devices operating at 3V

The data capture, start up and modulation board control functions are powered at 3V.

A8.1.1 ICM555 timer

[ICM7] The static supply current is 70μA; therefore the static power dissipation:

$$P_{STAT} = I\ V = (7 \times 10^{-5})\ (3) = 2.1 \times 10^{-4}\ W$$

To approximate a worst case dynamic power consumption, capacitor C8 is assumed to require a *full* new charge on each cycle of the oscillator. Therefore, the dynamic power, P_{DCAP}, required for charging the capacitor is:

$$P_{DCAP} = \tfrac{1}{2}\ f_{OSC}\ C\ V^2 = \tfrac{1}{2}\ (3.2 \times 10^4)\ (1 \times 10^{-9})\ (3)^2 = 1.44 \times 10^{-4}\ W$$

The second dynamic power dissipation arises from the power supply crowbarring during CMOS transitions. Each crowbar lasts ~200ns, and has a worst case current flow of 2.5mA; the power dissipated, P_{DCROW} is:

$$P_{DCROW} = f_{OSC}\ I\ V\ t_{SPIKE} = (3.2 \times 10^4)(2.5 \times 10^{-3})(3)(2 \times 10^{-7}) = 4.8 \times 10^{-5}\ W$$

The 555 timer is active for 10ms; therefore the energy requirement is:

$$E_{TOT} = P\ t_{ACTIVE} = (2.1 \times 10^{-4} + 1.44 \times 10^{-4} + 4.8 \times 10^{-5})\ (1 \times 10^{-2}) = \mathbf{4.0\ \mu J}$$

A8.1.2 HC4040

[HC404] It has power dissipative capacitance of 20pF and operates with an input frequency of 32kHz. The output frequency is negligible; the static power dissipation is also negligible. The device is active for 10ms; therefore its energy requirement is:

$$E_{TOT} = (C_{PD} \, V^2 \, f_{IN}) \, t_{ON} = (2 \times 10^{-11}) \, (3)^2 \, (3.2 \times 10^4) \, (1 \times 10^{-2}) = 0.06 \mu J$$

Hence the power consumption of this device is neglected.

A8.1.3 PVD1354

[PVD1] The photodiode supply current is limited by resistor R9; the voltage drop over the photodiode is 1.2V. Therefore the current is limited to:

$$I = V/R = 1.8 / 680 = 2.65mA$$

The photodiode is active for 10ms; therefore, the energy requirement for this device is:

$$E_{TOT} = I \, V \, t_{ON} = (2.65 \times 10^{-3}) \, (3) \, (1 \times 10^{-2}) = \textbf{79.5} \mu \textbf{J}$$

A8.1.4 NMP0103 smoothing capacitance

[NMP0] The 3V power supply includes a total of 22μF of smoothing capacitance. The energy requirement for their charging once per operation is:

$$E_{TOT} = \frac{1}{2} \, C \, V^2 = \frac{1}{2} \, (22 \times 10^{-6}) \, (3)^2 = \textbf{99} \ \mu \textbf{J}$$

A8.1.5 Other capacitance

The sum of the other capacitance powered at 3V is 1μF. The energy required for their charging once per operation is:

$$E_{TOT} = \tfrac{1}{2}\, C\, V^2 = \tfrac{1}{2}\, (1 \times 10^{-6})\, (3)^2 = \textbf{4.5 μJ}$$

A8.1.6 Other devices

All other devices powered at 3V consume negligible power (as their input and output frequencies are negligible).

A8.1.7 Total energy requirement, 3V board

[NMP0] The NMP0103 device has an efficiency of 80%. Taking the worst case assumption (that all power provided to the 3V circuitry is derived from this device) the total power consumed by the 3V-supplied circuitry is:

$$E_{3VOLT,TOT} = ((4.5 + 99 + 79.5 + 4.0) \times 10^{-6})\, (100/80) = \textbf{234 μJ}$$

A8.2 Devices powered by the 5V supply

The modulation function and transmitter are powered at 5V.

A8.2.1 HA7210

[HA72] The static supply current for this device is 65μA; therefore its static power dissipation is:

$$P_{STAT} = I\, V = (6.5 \times 10^{-5})\, (5) = 3.25 \times 10^{-4}\ W$$

The dynamic power dissipation for this device is derived from the load capacitance; in the system 2 handset this is an OR-gate input and a HC4040 counter input (3.5pF each). Therefore:

$$P_{DYN} = f_{OSC} \, C_L \, V^2 = (1 \times 10^6) \, (7 \times 10^{-12}) \, (5)^2 = 1.75 \times 10^{-4} \text{ W}$$

This device is active for 9.5ms, so its energy requirement is:

$$E_{TOT} = ((1.75 + 3.25) \times 10^{-4}) \, (9.5 \times 10^{-3}) = \mathbf{4.8 \mu J}$$

A8.2.2 PVD1354

[PVD1] The current limiting resistor for the photodiode is R1 on board 2, 1kΩ. The photodiode voltage drop is 1.2V, therefore the current is limited to:

$$I = V / R = 3.8 \times 10^{-3} \text{ A}$$

The device is active for 4ms so its energy requirement is:

$$E_{TOT} = (3.8 \times 10^{-3}) \, (5) \, (4 \times 10^{-3}) = \mathbf{76.0 \mu J}$$

A8.2.3 RF9901 (whilst not transmitting)

[RF99t] The RF9901 transmitter uses 22mA whilst its output amplifier is disabled; it is in this state for 4ms so its energy requirement is:

$$E_{TOT} = (22 \times 10^{-3}) \, (5) \, (4 \times 10^{-3}) = \mathbf{440.0 \mu J}$$

A8.2.4 RF9901 (transmitting)

[RF99t] When the output amplifier is activated the RF9901 requires 28mA for the 40μs of transmission.

$$E_{TOT} = (28 \times 10^{-3}) (5) (40 \times 10^{-6}) = \mathbf{5.6 \mu J}$$

A8.2.5 MAX756 smoothing capacitance

The 5V output DC-DC converter requires a 10μF smoothing capacitor. The energy required to charge this capacitor (once per handset operation) is:

$$E_{TOT} = \frac{1}{2} C V^2 = \frac{1}{2} (10 \times 10^{-6}) (5)^2 = \mathbf{125.0 \mu J}$$

A8.2.6 RF9901 smoothing capacitance

The RF9901 requires a total of 11μF of smoothing capacitance. The energy required to charge this (once per handset operation) is:

$$E_{TOT} = \frac{1}{2} C V^2 = \frac{1}{2} (11 \times 10^{-6}) (5)^2 = \mathbf{137.5 \mu J}$$

A8.2.7 Other devices

No other devices have a significant energy requirement, as they are principally static, and only active for the brief 40μs transmission period.

A8.2.8 Total energy requirement, 5V board

[MAX7] The MAX756 operates with 85% efficiency. Therefore, the total energy consumed by the 5V devices is:

$$E_{5VOLT,TOT} = ((4.8 + 76 + 440 + 11.2 + 125 + 137.5) \times 10^{-6}) (100/85) = \mathbf{935 \mu J}$$

A8.3 Energy requirement for one operation of a system 2 handset

The grand total energy consumed by one handset operation is:

$$E = (234 + 935)\ \mu J = \mathbf{1163\ \mu J}$$

Appendix 9   OHPs presented during talks at UCL and Imperial

The following sections detail the questions set on overhead projections during the experiments at UCL and Imperial College. They are shown reduced in size; some were put up as overlays - the final OHP shown is included here. Before each talk the students were shown the following overhead to familiarise them with the purpose and use of the handsets. The originals were in colour.
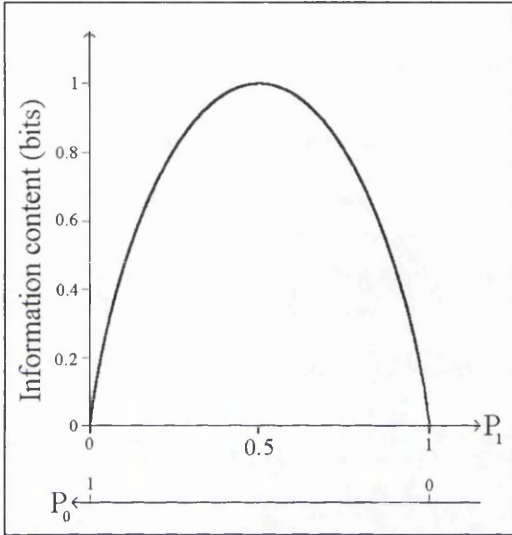

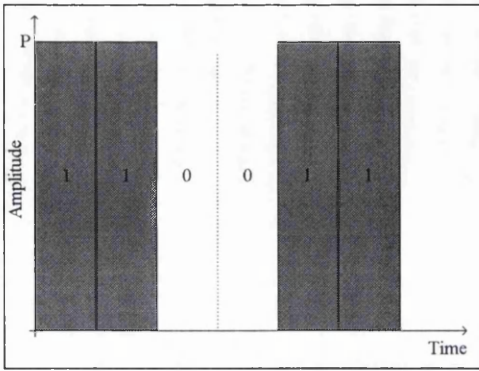When I ask you a question   At all other times
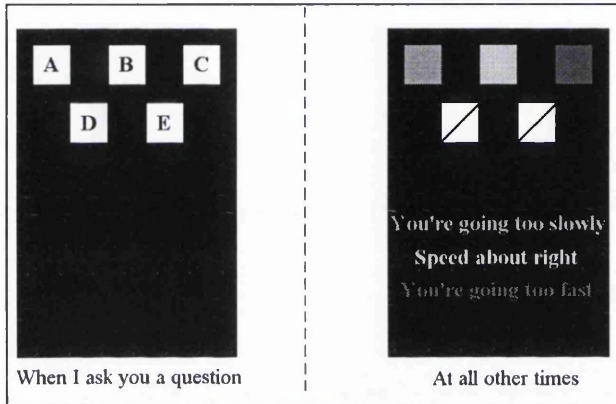
A9.1 Slides from talks at University College
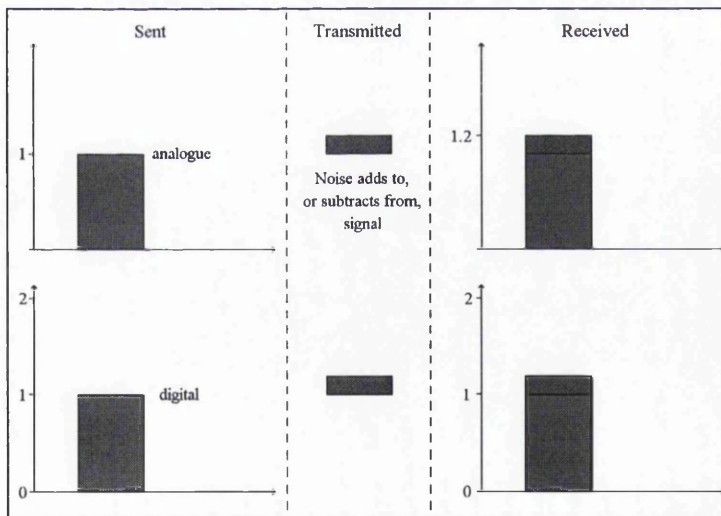
Which of the following sequences of written symbols, representing the number eight, has the most redundancy?
A.   Ag   (Afrikaans)
B.   1000   (Binary)
C.   BBDEB   (Arbitrary code)
D.   VIII   (Latin)
E.   8   (Decimal)

Which of the following *cannot* be used as a means to transmit more than a single piece of information between two points?

- A. Two yoghurt pots connected by a taut string
- B. Hitting logs with a stick
- C. A full moon in a specified lunar month
- D. Marbles
- E. A mirror that can both reflect and not reflect the sun

Which of the following bit sequences transmits the most information?

- A. 100001
- B. 101101
- C. 111000
- D. 110011
- E. None of the above

P

Amplitude

P/2

0

1

0

-P/2

0

1

0

Power $\propto P^2$

Power $\propto \dfrac{P^2}{2}$

Phase

μ

$\sqrt{2}\,\mu$

μ

0

μ
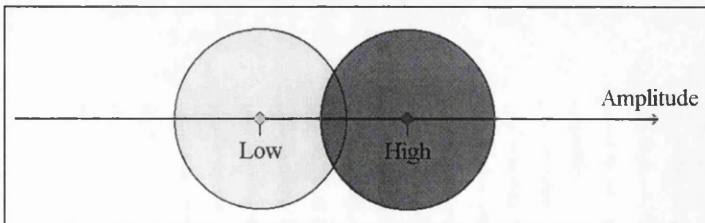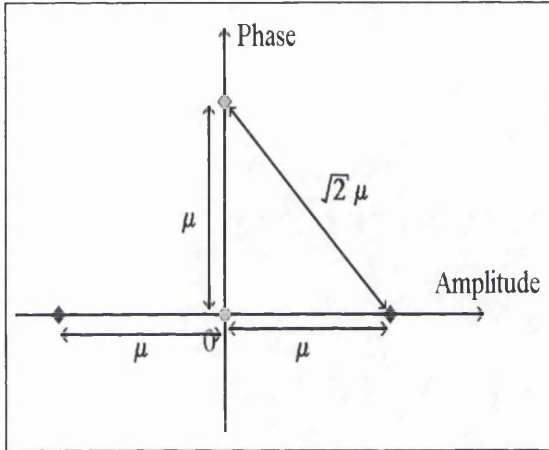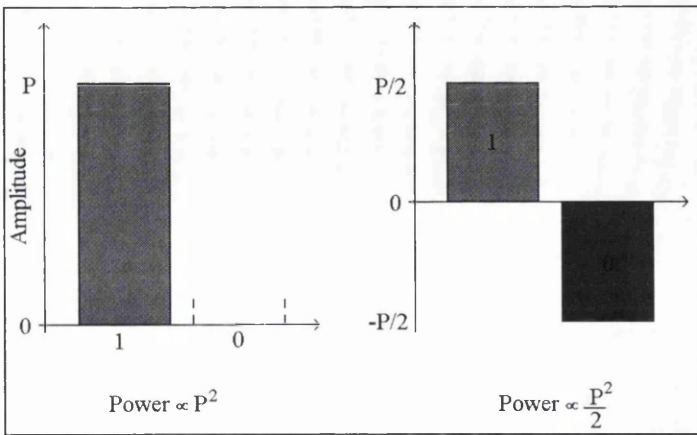
Amplitude

I wish to transmit a three level signal (with all levels equally likely), using DC voltage levels via a wire. The signals must have a separation greater than or equal to 4 volts. What amplitudes should my signals be to ensure minimum power consumption?

    A.   -4, 0 and 4 volts
    B.   0, 4 and 8 volts
    C.   0, 5 and 10 volts
    D.   -2, 0 and 2 volts
    E.   -6, -4 and 0 volts

P

Amplitude

2/3 P

1/3 P

11

10

01

00

Time

183

Signal with m levels
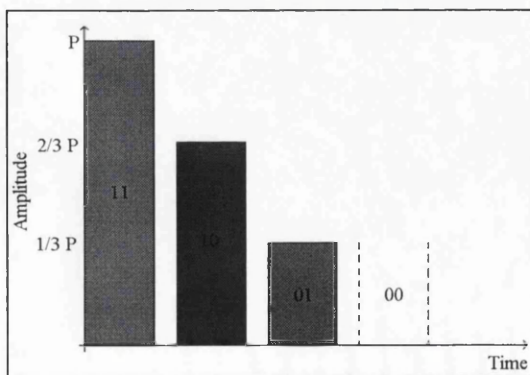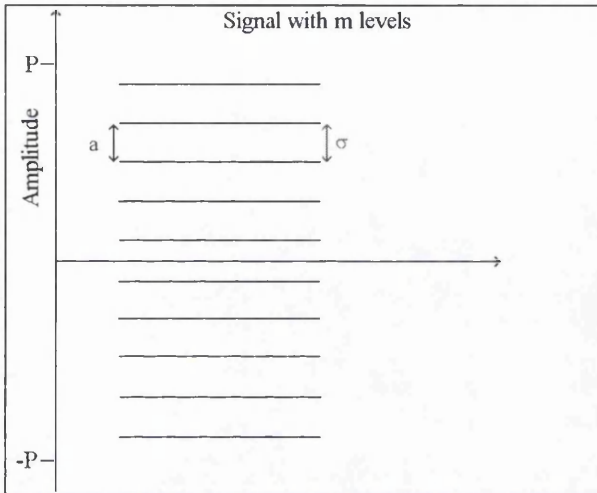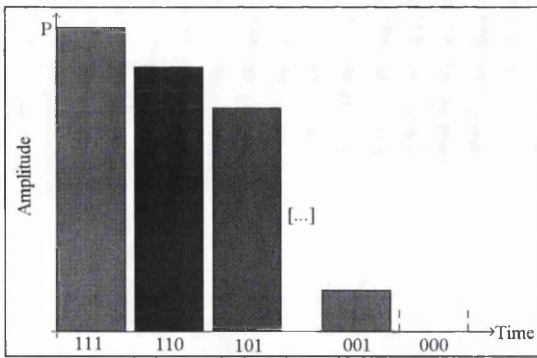
Dr. and Dr. Nerd live by an amplifier, and communicate in binary code using sound volume. The amplifier produces a noise which adds an root-mean-squared average of 50 dBm to every sound. What two sound volumes, used to transmit a one and a zero, will allow the Drs. Nerd to communicate?

A. 15 and 55 dBm

B. 50 and 100 dBm

C. 5 and 30 dBm

D. 20 and 80 dBm

E. 10 and 30 dBm

(note: decibels are ratios. A dBm is a decibel with respect to a microvolt, and is an absolute measure of the sound's volume)

τ = length of single bit

m = number of levels of signal

Number combinations in T seconds = $m^{T/\tau}$

Information transmitted $\propto (T/\tau)\log m$

Setting the logarithm to base 2:

Information = $(T/\tau) \log_2 m$

Capacity = Information/Time = $(1/\tau) \log_2 m$

$C = (1/\tau) \log_2 m$

But 1/time is a frequency!

If we have a range of frequencies over which we transmit, a bandwidth, of **W** Hertz, we can ideally fit in **2W** symbols sec^{-1}, giving:

$C = 2W \log_2 m$ or

$C = W \log_2 m^2$

Signal amplitude: $\pm a/2, \pm 3a/2, \dots \pm(m-1)a/2$

Signal power $\propto$ amplitude^2

average signal power $A = a^2 \left(\dfrac{m^2 - 1}{12} \right)$

So: $\quad C = W \log_2 m^2$ $\qquad$ becomes $\qquad C = W \log_2\left(1 + \dfrac{12A}{a^2}\right)$

But our signal separation, a is affected by noise:

$\quad a = K \sigma \qquad\qquad (K > 1)$
$\quad a^2 = K^2 B \qquad\qquad (\sigma^2 = B)$

Fitting this in to our equation we find that:

$$C = W \log_2\left(1 + \frac{12}{K^2}\frac{A}{B}\right)$$

Shannon's Law: $\qquad C = W \log_2\left(1 + \dfrac{A}{B}\right)$

Gives the maximum capacity that any given link can sustain **with no errors.**
As fundamental as the conservation of momentum or energy!

The Drs. Nerd have become depressed at their inability to communicate and are considering installing two-way radio for their binary communication. Which one has the greatest capacity?

- A. Bandwidth 60kHz, signal 1.5 volts, noise level 0.5 volts
- B. Bandwidth 20kHz, signal level 12 volts, noise level 2 volts
- C. Bandwidth 40kHz, signal level 9 volts, noise level 3 volts
- D. Bandwidth 20kHz, signal level 3 volts, noise level 0.5 volt
- E. Bandwidth 40kHz, signal level 6 volts, noise level 2 volts

---

Two signals are transmitted, one of -2 volts and one of +2 volts. During transmission, the noise function at the bottom of the page is added. What is the probability of being able to successfully distinguish the signals on reception?

- A. 15%
- B. 70%
- C. 30%
- D. 60%
- E. 40%

Percentage of the time the absolute magnitude of the noise does not exceed voltage value vs. Voltage

Number of modulation axes

16-states constellation for 9600 bits/sec modem

Which of the following signal constellations gives the best noise rejection?
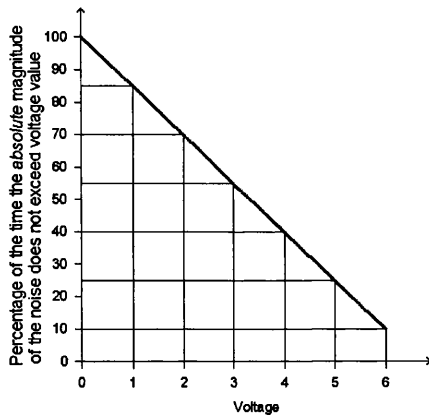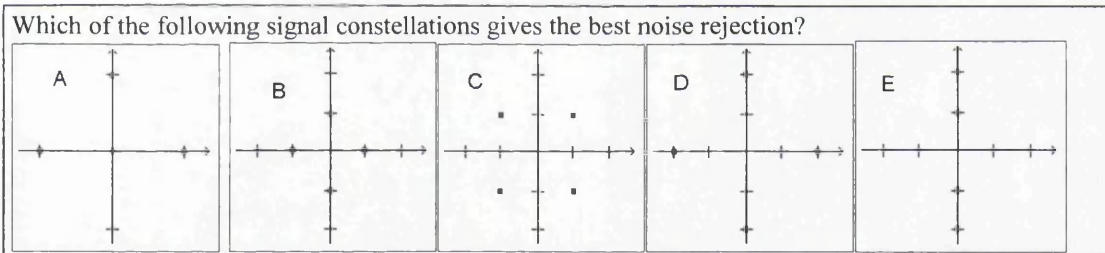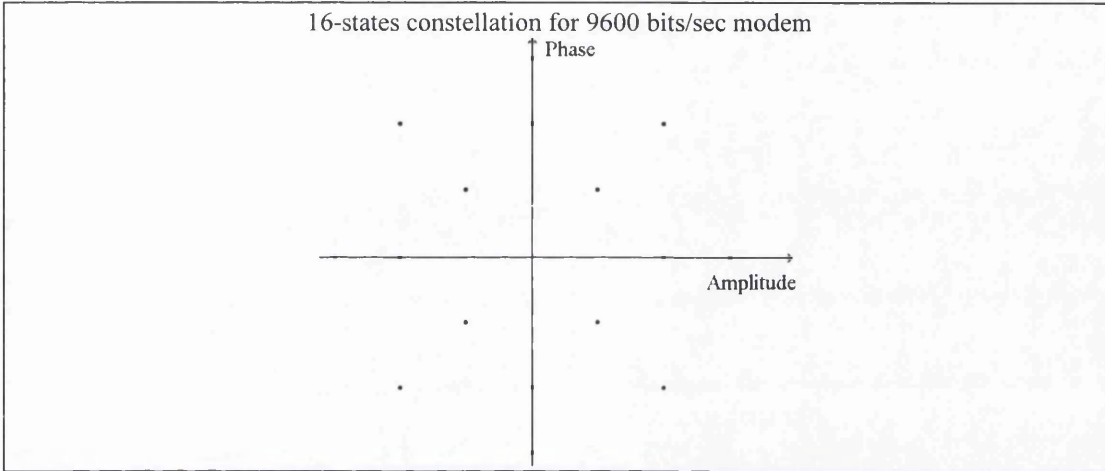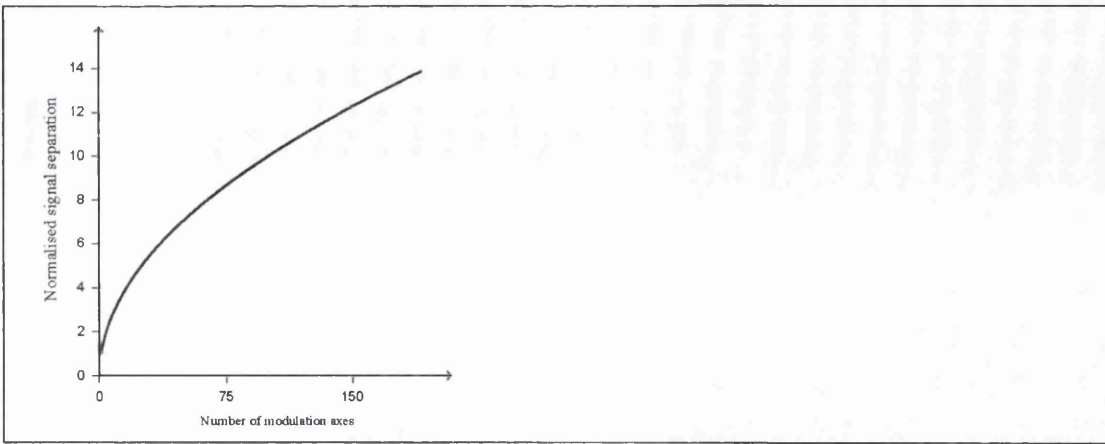
A9.2 Slides from Imperial College talks

The layers of the atmosphere
Note non-linear vertical scale!

186

Relative intensity

λ_{max} Wavelength

Solar energy in   Terrestrial heat out

The atmosphere

Mechanical energy   Heat

Work
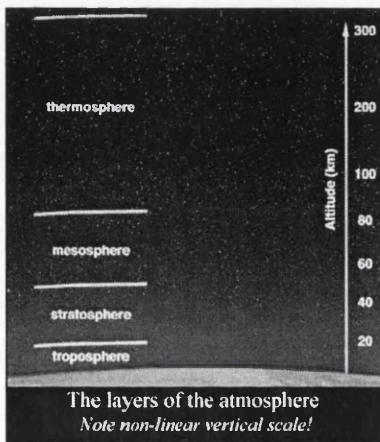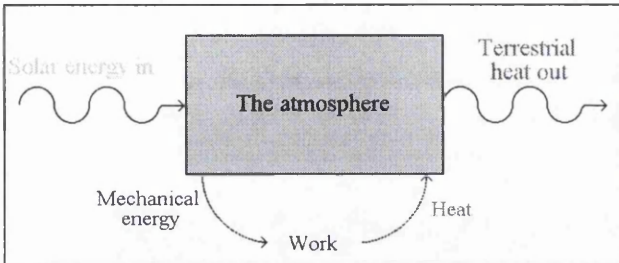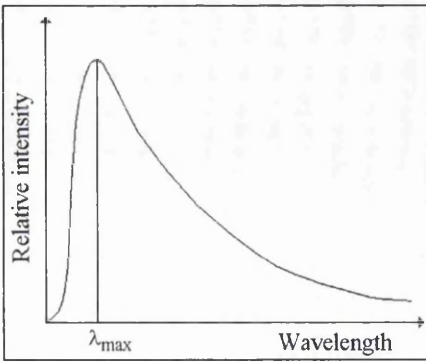
Which is the correct sequence of atmospheric layers, in order of altitude (lowest first)?
A.   Troposphere, stratosphere, thermosphere, mesosphere
B.   Troposphere, stratosphere, mesosphere, thermosphere
C.   Thermosphere, stratosphere, mesosphere, troposphere
D.   Thermosphere, mesosphere, stratosphere, troposphere
E.   Stratosphere, troposphere, mesosphere, thermosphere
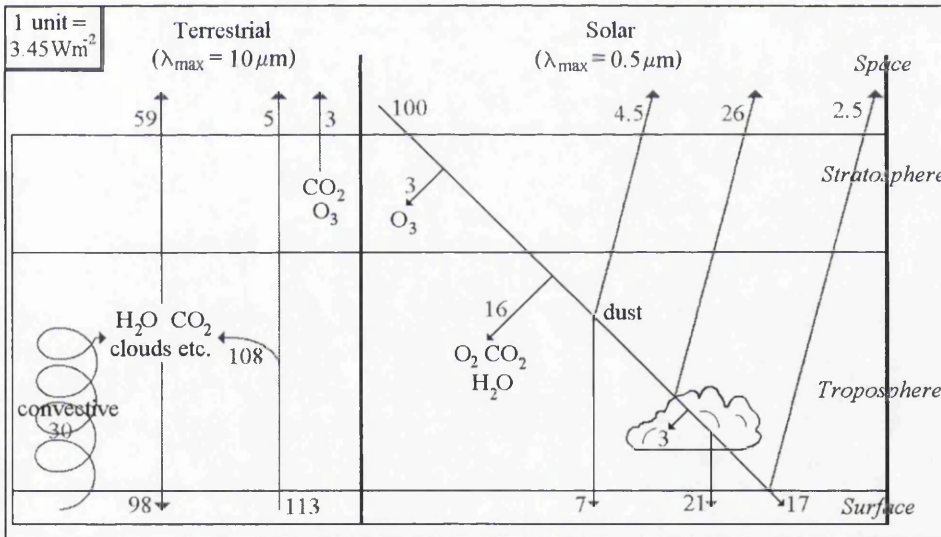
1 unit = $3.45 Wm^{-2}$

Terrestrial ($\lambda_{max} = 10 \mu m$)   Solar ($\lambda_{max} = 0.5 \mu m$)   Space

59   5   3   100   4.5   26   2.5

Stratosphere

CO_2 O_3 3 O_3

H_2O CO_2 clouds etc.   108   16   dust   Troposphere

O_2 CO_2 H_2O

convective 30   3

98   113   7   21   17   Surface

Why is the cloudless sky blue?
A.   Oxygen (O_2) is slightly blue
B.   The sky isn't blue - this is an optical illusion
C.   Because of differences in the scattering of light
D.   Because clouds leave a blue residue
E.   Because it's a reflection of the sea

187
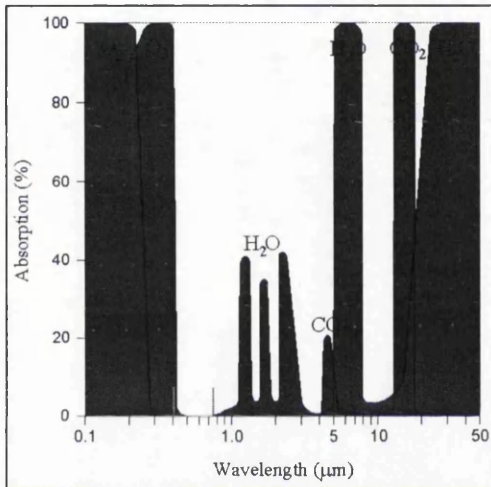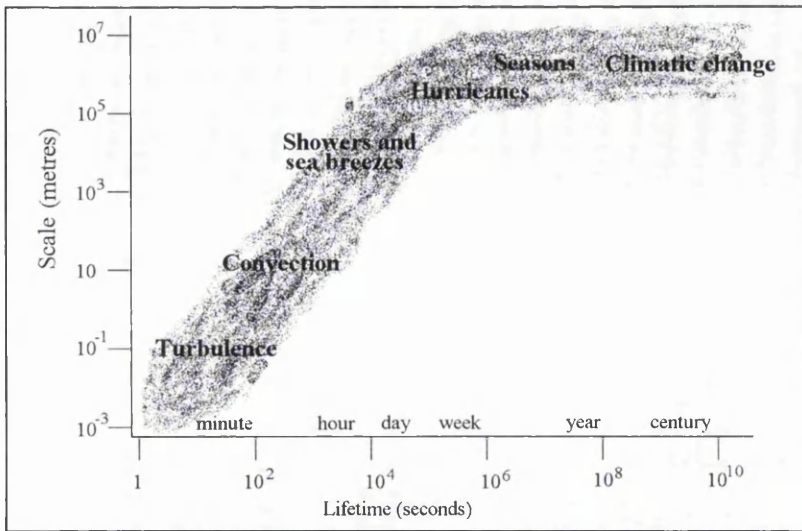
Which gas has causes the greatest proportion of the greenhouse effect in the Earth's atmosphere?

A. H_2O (water vapour)
B. N_2 (nitrogen)
C. CO_2 (Carbon dioxide)
D. CH_4 (methane)
E. O_3 (ozone)

Why is there so much fuss about artificially produced carbon dioxide (CO_2) emissions then?

A.   The carbon dioxide is additional and interferes in a complex set of dynamic equilibria
B.   It's purely political in motivation
C. CO_2 is often emitted at higher than ambient   temperatures
D.   The carbon dioxide (CO_2) is produced at ground level and causes great problems as it rises
E.   It's a different type of carbon dioxide
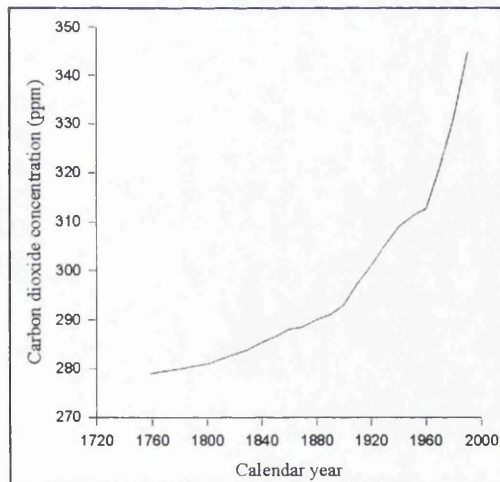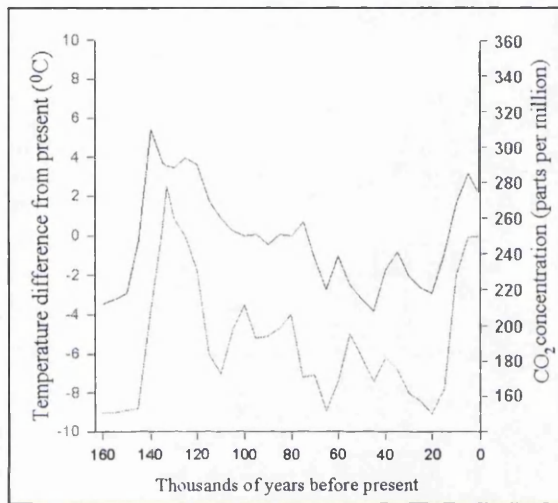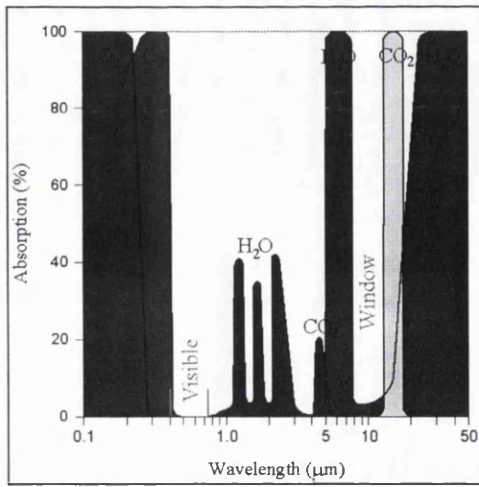
The mean of current estimates of the effects of human emissions predict a temperature rise of some 2-5 °C for a doubling of atmospheric CO_2. What percentage is that of the naturally occurring greenhouse effect?

A.   6 - 14%
B.   2 - 5%
C.   100 - 250%
D.   0.9 - 2.25%
E.   200 - 500%

If BSE leads to reduced cattle stocks, how will this influence the greenhouse effect?
A.   It will lead to a runaway greenhouse effect with the Earth's temperature rising by 500 °C
B.   It will lead to a lesser greenhouse effect
C.   It will lead to a greater greenhouse effect
D.   It will have no effect
E.   It will completely remove the greenhouse effect, causing a temperature drop of 35 °C
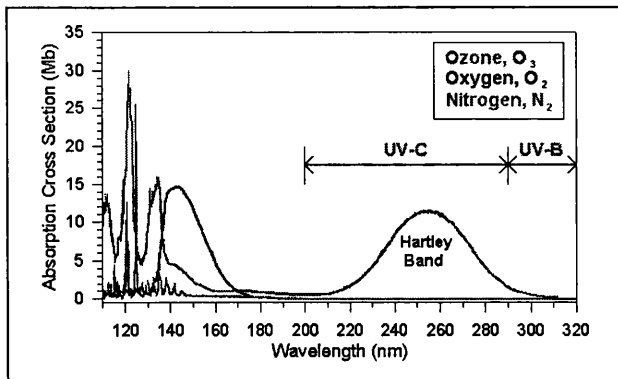
Which of the following is NOT a likely effect of increased UV intensity at the Earth's surface?
A.   Increased occurrence of cataracts in humans and animals
B.   Changes in the chemical composition of the troposphere
C.   Increased occurrence of skin cancers in humans and animals
D.   Accelerated degradation of synthetic polymer materials
E.   The sun will appear to be brighter to humans

Ozone is formed by photodissociation:

$$O_2 + photon \Rightarrow O_{radical} + O_{radical}$$
$$O_{radical} + O_2 \Rightarrow O_3$$
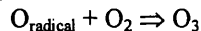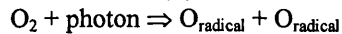
But! Chlorine catalyses its destruction *and* interferes with its production, for example:

$$Cl + O_3 \Rightarrow ClO + O_2 \qquad\qquad ClO + O \Rightarrow Cl + O_2$$
$$Net:\ Cl + O_3 + O \Rightarrow 2O_2 + Cl$$
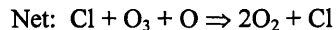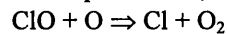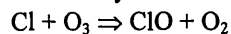
During darkness it forms temporary sink compounds (e.g. $OClO$ and Cl_2O) that photodissociate

Where is the "hole" in the ozone layer?
A.   Above the North pole
B.   Above the Equator
C.   Above the South pole
D.   Above both poles
E.   Above both poles and the equator

190

Appendix 10   Product-moment correlation of experimental data

The product moment correlation coefficient, r, between two variables, x and y, is conveniently defined by:

$$r = \left(\frac{(\overline{XY}) - (\overline{X})(\overline{Y})}{\sigma_x \sigma_y} \right) \left(\frac{N}{(N-1)} \right) \qquad \text{Eq. A10.1}$$

where $\overline{X}$ is the mean value of x, $\overline{Y}$ the mean of y, σ_x and σ_y the standard deviations of the x and y respectively, $\overline{XY}$ the mean of the *product* of x and y and N is the number of sample points.[1] The correlation coefficient determines whether there is a linear relation between the two variables - it will be 1 for perfectly aligned samples on a line with positive gradient, -1 for perfectly aligned sample on a negative gradient, and 0 for a sample exhibiting no correlation. It can only be used to test for a linear correlation where there are no outlying points and where the sample is a randomly selected fraction of the larger population. One can compare the data to various significance tests; the one used in this data is the twin-tailed significance test (that is, it tests the *modulus* of the correlation coefficient ignoring whether the correlation is positive or negative) [SIEG]. There are four levels of confidence in the significance of the correlation coefficient: 90% indicating some confidence, 95% indicating high confidence; 99% indicating very high confidence and 99.9% indicating extremely high confidence.

A10.1 Engagement

This data has as its x variable the percentage of questions answered correctly during talks, and its y variable is the total number of times students pressed a button to express an opinion on the speed of delivery. The data from the UCL experiments are shown in table A10.1; the data from the Imperial College experiments are shown in table A10.2

---

[1] one does, however, have to be careful of rounding errors whilst calculating the correlation coefficient using this equation

x	y	xy	Coefficients
0	0	0	$\overline{X}$ = 40.9091
12.5	2	25	$\overline{Y}$ = 8.3636
12.5	19	237.5	$\overline{XY}$ = 407.3864
12.5	9	112.5	σ_X = 21.8837
25	2	50	σ_y = 7.0545
25	0	0	N = 22
25	2	50	**r = 0.44**
37.5	0	0	
37.5	7	262.5	
37.5	8	300	
37.5	3	112.5	
37.5	5	187.5	
37.5	12	450	
50	3	150	
50	19	950	
50	24	1200	
62.5	4	250	
62.5	10	625	
62.5	10	625	
75	14	1050	
75	15	1125	
75	16	1200	

Table A10.1 -- Engagement, UCL students

This gives a value of r, 0.44, above the 95% confidence level which occurs at r=0.42 for 22 points [SIEG].

x	y	xy	Coefficients
0	1	0	$\overline{X} = 39.8438$
0	2	0	$\overline{Y} = 10.8125$
12.5	3	37.5	$\overline{XY} = 483.2031$
12.5	6	75	$\sigma_X = 22.0971$
12.5	12	150	$\sigma_y = 5.7217$
12.5	16	200	$N = 32$
25	2	50	**$r = 0.43$**
25	4	100	
25	9	225	
25	12	300	
25	17	425	
37.5	4	150	
37.5	6	225	
37.5	8	300	
37.5	9	337.5	
37.5	11	412.5	
37.5	11	412.5	
37.5	15	562.5	
37.5	18	675	
50	9	450	
50	15	750	
50	16	800	
50	17	850	
50	22	1100	
62.5	8	500	
62.5	7	437.5	
62.5	14	875	
62.5	15	937.5	
62.5	19	1187.5	
75	12	900	
75	19	1425	
87.5	7	612.5	

Table A10.2 -- Engagement, Imperial students

This value of r, 0.43, exceeds the 95% confidence level for 32 points, 0.35 [SIEG].

A10.2 Aptitude

This data has as its x variable the student's overall academic mark for the previous year, and its y variable is the percentage of questions answered correctly during talks. The data from the UCL experiments are shown in table A10.3; the data from the Imperial College experiments are shown in table A10.4.

x	y	xy	Coefficients
46.7	0	0	$\overline{X} = 63.1909$
61.9	25	1547.5	$\overline{Y} = 40.9091$
57.6	37.5	2160	$\overline{XY} = 2781.25$
62.3	25	1557.5	$\sigma_X = 12.7142$
61.3	25	1532.5	$\sigma_y = 21.8837$
78.3	50	3915	$N = 22$
58.9	12.5	736.25	$r = 0.74$
46	62.5	2875	
61	37.5	2287.5	
92	75	6900	
60.3	37.5	2261.25	
53	37.5	1987.5	
65.7	37.5	2463.75	
82.4	75	6180	
72.6	50	3630	
44.1	12.5	551.25	
46.1	12.5	576.25	
72.4	62.5	4525	
76	62.5	4750	
76.1	75	5707.5	
57	50	2850	
58.5	37.5	2193.75	

Table A10.3 -- Aptitude, UCL students

This value of r, 0.74, exceeds the 99.9% confidence level for 22 points, 0.65 [SIEG].

x	y	xy	Coefficients
47	0	0	$\overline{X} = 60.8125$
54	0	0	$\overline{Y} = 39.8438$
55	12.5	687.5	$\overline{XY} = 2605.469$
42	12.5	525	$\sigma_X = 10.6391$
46	12.5	575	$\sigma_y = 22.0971$
47	12.5	587.5	$N = 32$
52	25	1300	**$r = 0.80$**
60	25	1500	
50	25	1250	
56	25	1400	
62	25	1550	
51	37.5	1912.5	
55	37.5	2062.5	
54	37.5	2025	
61	37.5	2287.5	
64	37.5	2400	
75	37.5	2812.5	
64	37.5	2400	
55	37.5	2062.5	
57	50	2850	
63	50	3150	
71	50	3550	
69	50	3450	
59	50	2950	
61	62.5	3812.5	
71	62.5	4437.5	
79	62.5	4937.5	
65	62.5	4062.5	
60	62.5	3750	
76	75	5700	
84	75	6300	
81	87.5	7087.5	

Table A10.4 -- Aptitude, Imperial students

This value of r, 0.81, exceeds the 99.9% confidence level for 32 points, 0.56 [SIEG].

A10.3 Self-satisfaction

This data has as its x variable the percentage of student prompted button presses that indicated dissatisfaction - that were 'too fast' or 'too slow', and its y variable is the

195

percentage of questions answered correctly during talks. The data from the UCL experiments are shown at table A10.5; the data from the Imperial College experiments are shown in table A10.6.

x	y	xy	Coefficients
50	25	1250	$\overline{X} = 68.5789$
100	25	2500	$\overline{Y} = 44.0789$
33.3	50	1665	$\overline{XY} = 2777.697$
100	12.5	1250	$\sigma_X = 23.3793$
50	62.5	3125	$\sigma_y = 20.9863$
85.7	37.5	3213.75	N = 19
71.4	75	5355	**r = -0.53**
100	37.5	3750	
100	37.5	3750	
78.9	50	3945	
80	37.5	3000	
26.7	75	2002.5	
75	50	3750	
58.9	12.5	736.25	
83.3	37.5	3123.75	
66	12.5	825	
50	62.5	3125	
43.8	75	3285	
50	62.5	3125	

Table A10.5 -- Self-satisfaction, UCL students

This value of r, 0.53, exceeds the 95% confidence level for 19 points, 0.46 [SIEG].

x	y	xy	Coefficients
100	0	0	$\overline{X} = 76.8438$
100	0	0	$\overline{Y} = 39.8438$
100	12.5	1250	$\overline{XY} = 2969.531$
50	12.5	625	$\sigma_X = 18.2458$
67	12.5	837.5	$\sigma_y = 22.0971$
94	12.5	1175	$N = 32$
100	25	2500	**$r = -0.24$**
100	25	2500	
44	25	1100	
50	25	1250	
82	25	2050	
75	37.5	2812.5	
83	37.5	3112.5	
75	37.5	2812.5	
56	37.5	2100	
100	37.5	3750	
64	37.5	2400	
87	37.5	3262.5	
61	37.5	2287.5	
89	50	4450	
100	50	5000	
69	50	3450	
77	50	3850	
64	50	3200	
63	62.5	3937.5	
100	62.5	6250	
71	62.5	4437.5	
60	62.5	3750	
84	62.5	5250	
50	75	3750	
58	75	4350	
86	87.5	7525	

Table A10.6 -- Self-satisfaction, Imperial students

This value of r, 0.24, does not exceed 90% confidence level for 32 points, 0.30 [SIEG].

Appendix 11 Parts list, system 2, board 1

The parts that are referred to by letter on figure 8.2 are listed in table A11.1. The resistors R7 and R8 should be within 1% of their nominal value.

R1a to R1e	all 4.7kΩ	C1a to C1e	all 0.1µF
R2a to R2e	all 100kΩ	C2a to C2e	all 0.1µF
R3a to R3e	all 47kΩ	C3a to C3e	all 0.1µF
R4	4.7kΩ	C4	22nF
R5	2.2kΩ	C5	0.1µF
R6	20kΩ	C6	220nF
R7 (1% tol.)	14kΩ	C7	0.1µF
R8 (1% tol.)	16.9kΩ	C8	1nF
R9	680Ω	C9	1nF
R10	4.7kΩ	C10	0.1µF
R11	47kΩ	C11	10µF
R12	10kΩ	C12	0.1µF
R13	1kΩ		
R14	1MΩ		
D1	1N5711 (or equivalent)		
D2	1N5817 (or equivalent)		
Junction transistors	2N5089 (BC549C when stocks are exhausted)		
FET	2N7000(SLX)		
L1	22µH		

Table A11.1 -- The component values for board 1, system 2 (figure 8.2)

Appendix 12   Papers published

The following papers have been published by the author during this research:

[GILB] Gilbert, M., Ash, E. A., "Electrifying lectures: finding a role for students", *Proceedings of the 14th International Conference on Technology and Education*, Vol. 1, Oslo, August 1997, pp.291-293

Poulis, J., Massen, C., Robens, E., Gilbert, M., "Physics lecturing with Audience Paced Feedback", *American Journal of Physics*, Vol. 66, No. 5, 1998, pp.439-441

References

[ABRA] Abramson, N., "The Aloha System", in Kuo, F., ed., *Computer-Communication Networks*, Prentice-Hall, 1973, pp.501-516

[ADAI] Adair, G., "The Hawthorne effect: a reconsideration of the methodological Artefact", *Journal of Applied Psychology*, Vol. 69, No. 2, 1984, pp.334-345

[ALIF] Aliftiras, G., "Receiver Implementations for a CDMA Cellular System", MSc. Thesis, Virginia Polytechnic Institute and State University, July 1996, pp.40

[APEM] Apem Components Ltd., Drakes Drive, Long Crendon, Buckinghamshire, HP18 9BA

[AM72] *CMOS FIFO Memory Products Data Book/Handbook*, Advanced Micro Devices Inc., 1992, pp.2-19

[BLAH] Blahut, R. E., *Theory and Practice of Error Control Codes*, Addision-Wesley, 1984

[BLIG] Bligh, D. A., *What's the Use of Lectures*, 2nd. Ed., published by D. A. and B. Bligh, New Barnet, 1972

[BOSW] Boswell, J., (annotations Malone, E.), *The Life of Samuel Johnson, LL.D.*, 6th Ed. reprint, George Routledge and Sons (London), p.412; first published in 1791

[BOTT] Bottomley, G.E., "Improved successive cancellation of DS-CDMA signals using signal orthogonalization", *Proc. IEEE International Conference on Universal Personal Communications*, 1996, pp.141-144

[BRUN] Bruner, J., *Toward a Theory of Instruction*, Harvard University Press, 1966

[BURT] Burton, E., "*The whole idea: Gestalt psychology*", in Burton, E., ed., *Thinking in Perspective*, Methuen, 1978

[BYLA] Bylanski, P., & Ingram, D. G. W., *Digital Transmission Systems*, Peter Pereginus Ltd., 1976

[COIL] Coilcraft UK, 21 Napier Place, Wardpark North, Cumbernauld, Scotland, G68 0LL

[COMS] TDMA: Part I, Special Issue, *Comsat Technical Review*, vol. 15, No. 2b, pp.361-525, Autumn 1985 and TDMA: Part II, Special Issue, *Comsat Technical Review*, vol. 16, no. 1, pp.1-298, Spring 1986

[DAVI] Davies, N., et al, *Information Technology Assisted Teaching and Learning in UK Higher Education*, Consortium of Telematics for Education, University of Exeter, January 1997, Recommendation 4

[DEAR] Dearing, R., et al, *"Higher Education in the Learning Society"* (National Committee of Inquiry into Higher Education), HMSO, 1997

[DIGI] Digi-Key Corporation, 701 Brooks Avenue South, PO Box 677, Thief River Falls, MN, USA

[DG41] "DG417/418/419 Precision CMOS Analog Switches", S-52880-Rev. D, Temic Semiconductors, April 1997

[DIXO] Dixon, R. C., *Spread Spectrum Systems*, 2nd Ed., John Wiley & Sons, 1984

[DRAP] Draper, S., "Niche-based success in CAL", *Proceedings of Computer Aided Learning 97*, Exeter, March 1997, Paper No. 95a

[DURA] *Comprehensive Battery Guide*, Duracell Incorporated, 1998

[ECONa] *The knowledge factory, A Survey of Universities,* The Economist, October 4 1997, p.4

[ECONb] "Under examination" and "Fee fight", *The Economist*, Vol. 351, No. 8122, 5th June 1999, pp.64-66

[ELTO] Elton, L., *Teaching in Higher Education*, Kogan Page, 1987, p.168

[FANF] Fan, F. and Darnell, M., *Sequence Design for communication applications*, Research Studies Press, 1996

[FARN] *Electronic Components*, Premier Farnell PLC, Canal Road, Leeds, LS12 2TU

[FRER] Frerking, M. E., *Digital Signal Processing in Communication Systems*, Van Nostrand Reingold, 1994, p.399

[GATE] Gates, W., "Linked up for learning: Using technology on campus", *Educational Record*, Vol. 77, No. 4, 1994, pp.34-41

[GIBB] Gibbs, G., Jenkins, A., "An introduction: the context of changes in class size", in *Teaching Large Classes in Higher Education: How to Maintain Quality with Reduced Resources*, Gibbs, G., Jenkins, A., Eds., Kogan Page, 1992, pp.11-18

[GINS] Ginsbury, H. P., Opper, S., eds., *Piaget's Theory of Intellectual Development*, Prentice-Hall, 1969

[GNEN] Gnendenko, B. V., Kovalenko, I. N., *Introduction to Queueing Theory*, S. Monson (Jerusalem), 1968, pp.1-18

[GOLD] Gold, R., "Optimal Binary Sequences for Spread Spectrum Multiplexing", *IEEE Transactions on Information Theory*, October 1967, pp.619-621

[GOLO] Golomb, S. W., Ed., *Digital Communications with Space Applications*, Prentice-Hall, 1964, pp.135-160

[GRAY] Gray, R. M., Davisson, L. D., *Random Processes: A Mathematical Approach for Engineers*, Prentice-Hall, 1986

[GREG] Gregory, I., "A New Look at the Lecture Method", *British Journal of Educational Technology*, Vol. 6, No. 1, 1975, pp.55-62

[HA72] *HA7210, HA7211 Data Sheet*, File Number 3389.6, Harris Semiconductor, November 1996

[HAMB] Hambley, A. R., *An Introduction to Communication Systems*, W. H. Freeman and Co., 1990

[HC16] "8-bit parallel-in/serial-out shift register 74HC/HCT165", Philips Semiconductors, December 1990

[HC404] "12-stage binary ripple counter 74HC/HCT4040", Philips Semiconductors, December 1990

[HC406] "74HC/HCT4060 14-stage binary ripple counter with oscillator", Philips Semiconductors, December 1990

[HC74] "Dual D-type flip-flip with set and reset; positive-edge trigger 74HC/HCT74", Philips Semiconductors, February 1998

[HCMO] *Family Specifications: HCMOS family characteristics*, File IC06, Philips Semiconductors, 1988

[HOLL] Holling, K., "Feedback Classrooms", *Industrial Training International*,
Vol.2, No.6, 1967, pp.254-256

[HORO] Horowitz, P., Hill, W., *The Art of Electronics*, 2nd. Ed., Cambridge
University Press, 1994

[HT12e] "2^{12} Series of Encoders", Holtek Semiconductor Inc., October 1997

[HT12d] "2^{12} Series of Decoders", Holtek Semiconductor Inc., October 1997

[ICM7] "ICM7555, ICM7556 General Purpose Timers", Harris Semiconductor
Corporation, File Number 2867.2, November 1996

[INTE] Intersolar Limited, Cock Lane, High Wycombe, Buckinghamshire, HP13
7DE

[IVES] Ives, J., "University TV Research at Leeds", *Journal of Educational
Technology*, Vol. 1, No. 2, 1970, pp.139-147

[JACK] Jackson, M., Prosser, M. T., "Less Lecturing, More Learning", *Studies in
Higher Education*, vol. 14, No. 1, 1989, pp.55-61

[KLEI] Kleinrock, K., *Queueing Systems Volume I: Theory*, John Wiley & Sons,
1975, pp.48-79

[LAMS] Lam, S. S., "A Carrier Sense Multiple Access Protocol for Local
Networks", *Computer Networks*, vol. 4, No. 1, January 1980, pp.21-32

[LASN] Lasnier, F., Ang, T. G., *Photovoltaic Engineering Handbook*, Adam Hilger
(New York), 1990

[LAUR] Laurillard, D., *Rethinking University Teaching*, Routledge, 1993

[LAYA] Layard, R., Oatey, M., "The Cost-Effectiveness of the New Media in
Higher Education", *British Journal of Educational Technology*, Vol. 4, No. 3, 1973,
pp.158-176

[LESG] Lesgold, A. M., ed., *Cognitive Psychology and Instruction*, Plenum Press,
1978

[LIBEa] Liber, O., "Viewdata and the World Wide Web; Information or Communication?", *Proceedings of Computer Aided Learning 97*, Exeter, March 1997, Paper No. 187a

[LIBEb] Liber, O., "Viewdata - 20th Century Caxton?", *Bulletin of Environmental Education*, No. 197/8, Streetwork (London), 1987

[LUQU] Luque, A., *Solar Cells and Optics for Photovoltaic Concentration*, Adam Hilger (Philadelphia), 1989

[MANS] Mansuripur, M., *Introduction to Information Theory*, Prentice-Hall, 1987

[MART] Marton, F., Hounsell, D., eds., *The experience of learning*, Scottish Academic Press, 1984

[MATT] Matthews, M. R., *Constructivism in Science Education: A Philosophical Examination*, Kluwer Academic Publishers, 1998

[MAX7] "3.3V/5V/Adjustable-Output, Step-Up DC-DC Converters MAX756/MAX757", Maxim Integrated Products, Rev. 2, January 1995

[MAX7b] "-5V/-12V/-15V High-Efficiency Low-Q Inverting DC-DC converters", Maxim Integrated Products, Rev. 1, March 1994

[MAZU] Mazur, E., *Peer Instruction A User's Manual*, Prentice-Hall, 1997, p.14

[McKE] McKeachie, W. J., *Teaching Tips: A Guidebook for the Beginning College Teacher*, 8th Ed., Lexington, D. C. Heath and Co., 1986, p.69

[MICH] Michelson, A. M., Levesque, A. H., *Error-Control Techniques for Digital Communications*, Wiley, 1985

[MULG] Mulgrew, P. G., Thompson, J., *Digital signal processing: concepts and applications*, Macmillan, 1999

[NMP0] "Discontinued datasheet: NMP0103", Newport Components, October 1997

[PARS] Parsons, H. M., "What happened at Hawthorne?", *Science*, Vol. 183, 1974, pp.922-932

[PASU] Pasupathy, S., "Minimum Shift Keying: A Spectrally Efficient Modulation", *IEEE Communications Magazine*, vol. 17, No. 4, July 1979, pp.14-22

[PATE] Patel, P., Holtzman, J., "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system", *IEEE Journal on Selected Areas in Communications*, Vol. 12, No. 5, 1994, pp.796-807

[PAVL] Pavlov, I. P., translated by Anrep, G. V., *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*, Oxford University Press, 1927

[POPP] Popper, K. R., *Conjectures and refutations: the growth of scientific knowledge*, 5th ed. (revised), Routledge, 1992, 1989

[PORA] Porat, B., *A course in digital signal processing*, John Wiley and Sons, 1997

[POULa] Poulis, J., Massen, C., Robens, E., Gilbert, M., "Physics lecturing with Audience Paced Feedback", *American Journal of Physics*, Vol. 66, No. 5, 1998, pp.439-441

[POULb] Poulis, J., Massen, C., Robens, E., "Rationalisierung der Hochschulausbildung mit Audience Paced Feedback", *Das Hochschulwesen*, Vol. 41, 1993, p.91

[PROA] Proakis, J. G., Salehi, M., *Communication Systems Engineering*, Prentice-Hall, 1994

[PVD1] "Series PVD13 Microelectronic Power IC Relay", International Rectifier, PD1.024D

[RAED] Raedor Circuits, 7 Little Ridge, Ridgeway, Welwyn Garden City, Herts. AL7 2BH

[RF99r] "RF9902 FSK Receiver", Rev B2, RF Micro Devices, June 1996

[RF99t] "RF9901 FSK Transmitter", Rev D2, RF Micro Devices, December 1998

[ROES] Roesthlisberger, F. J., Dickson, W. J., *Management and the Worker*, Harvard University Press, 1946

[SARW] Sarwate, D. V. and Pursley, M. B., "Crosscorrelation properties of pseudorandom and related sequences", *Proceedings of the IEEE*, Vol. 68, No. 5, 1980, pp.593-620

[SCHW] Schwartz, M., *Information Transmission, Modulation, and Noise*, 4th Ed., McGraw-Hill, 1990

[SHAN] Shannon, C. E., "Communication in the Presence of Noise", *Proceedings of the IRE*, Vol. 37, January 1949, pp.10-21

[SIEG] Siegel, A. F., *Statistics and Data Analysis An Introduction*, John Wiley & Sons, 1988, pp.422-431

[SIL4] "SILRX Series UHF FM Superhet Receiver Module", Radio-Tech Limited, February 1997

[SKIN] Skinner, B. F., *Science and Human Behaviour*, Macmillan, 1953

[STEK] Steklova, V. A., "Matematicheskie metody teorii massovogo ovsluzhivania" ("Mathematical Methods in the Theory of Queueing"), *Trudy Matematicheskogo Instituta*, Vol. 49, 1955; the English translation is in *Griffin's Statistical Monographs and Courses*, No. 7, London, 1960

[STRE] Stremler, F. G., *Introduction to Communication Systems*, 3rd Ed., Addison-Wesley, 1990

[TAKA] Takahashi, K., Kongai, M., translated by Apps, F. R. D., *Amorphous silicon solar cells*, North Oxford Academic (London), 1986, p.92

[TAPL] Taplin, G., "The Cosford Cube: a simplified form of student feedback", *Industrial Training International*, Vol. 4, No. 5, 1969, pp.18-21

[TAUB] Taub, H., and Schilling, D. L., *Principles of Communication Systems*, 2nd. Ed., McGraw-Hill Inc, 1986

[TEKO] Teko S.p.A., Via dell'Industria, 7 40068, S. Lazzaro di Savena, Bologne, Italy

[THOR] Thorndike, E. L., *Education A First Book*, MacMillan (New York), 1912

[TIRK] Tirkel, A. Z., "Cross-Correlation of M-Sequences - Some Unusual Coincidences", *IEEE International Symposium on Spread Spectrum Techniques and Applications*, Mainz Germany, 1996, pp.969-973

[TREB] Treble, F. C., Ed., *Generating Electricity from the Sun*, Pergamon Press, 1991

[TSIV] Tsividis, Y. P., *Operation and Modelling of the MOS Transistor*, McGraw-Hill International Editions, 1988

[TXM4] "Low Power UHF Radio Telemetry Module UK Version - TXM-418-F", Radio-Tech Limited, February 1997

[VERDa] Verdú, S., "Multiuser detection", in *Advances in Detection and Estimation*, JAI Press, 1993

[VERDb] Verdú, S., "Minimum probability of error for asynchronous gaussian multiple-access channels", *IEEE Transactions on Information Theory*, Vol. IT-32, January 1986, pp.85-96

[VITEa] Viterbi, A. J., *CDMA: principles of spread spectrum communication*, Addison-Wesley, 1995, pp.34-36

[VITEb] Viterbi, A. J., *Principles of Coherent Communication*, McGraw-Hill, 1966, figure 8.3

[WARD] Ward, A., Jenkins, A., "The problems of learning and teaching in large classes", in Gibbs, G., Jenkins, A., eds., op. cit., p.25

[WATS] Watson, W. B., *Behaviourism*, (revised edition), University of Chicago Press, 1930 (seventh impression 1967)

[WEBB] Webb, C., "Feedback Teaching Systems", *Industrial Training International*, Vol.5, No.5, 1970, pp.243-245

[ZEIM] Zeimer, R. E., Tranter, W. H., *Principles of Communications Systems, Modulation and Noise*, 4th Ed., Houghton Mifflin, 1995