Chapter 10
Writing a Talk

My recommendations amount to this ...
Make your lecture simple (special and concrete);
be sure to prove *something and* ask *something;*
prepare, in detail;
organize the content and adjust to the level of the audience;
keep it short, and, to be sure of doing so,
prepare it so as to make it flexible.

— PAUL R. HALMOS, *How to Talk Mathematics* (1974)

I always find myself obliged,
if my argument is of the least importance,
to draw up a plan of it on paper and
fill in the parts by recalling them to mind,
either by association or otherwise.

— MICHAEL FARADAY[16]

An awful slide is one which contains approximately a million numbers
(and we've left our opera glasses behind).
An awful lecture slide is one which shows a complete set of
engineering drawings and specifications for a super-tanker.

— KODAK LIMITED, *Let's Stamp Out Awful Lecture Slides* (1979)

[16] Quoted in [271, p. 98].

10.1. What Is a Talk?

In this chapter I discuss how to write a mathematical talk. By talk I mean a formal presentation that is prepared in advance, such as a departmental seminar or a conference talk, but not one of a series of lectures to students. In most talks the speaker writes on a blackboard or displays pre-written transparencies on an overhead projector. (From here on I will refer to transparencies as slides, since this term is frequently used and is easier to write and say.) I will restrict my attention to slides, which are the medium of choice for most speakers at conferences, but much of what follows is applicable to the blackboard. I will assume that the speaker uses the slides as a guide and speaks freely. Reading a talk word for word from the slides should be avoided; one of the few situations where it may be necessary is if you have to give a talk in a foreign language with which you are unfamiliar (Kenny [149] offers some advice on how to do this).

A talk has several advantages over a written paper [50].

1. Understanding can be conveyed in ways that would be considered too simplified or lacking in rigour for a journal paper.

2. Unfinished work, or negative results that might never be published, can be described.

3. Views based on personal experience are particularly effective in a talk.

4. Ideas, predictions and conjectures that you would hesitate to commit to paper can be explained and useful feedback obtained from the audience.

5. A talk is unique to you—no one else could give it in exactly the same way. A talk carries your personal stamp more strongly than a paper.

Given these advantages, and the way in which written information is communicated in a talk, it is not surprising that writing slides differs from writing a paper in several respects.

1. Usually, less material can be covered in a talk than in a corresponding paper, and fewer details need to be given.

2. Particular care must be taken to explain and reinforce meaning, notation and direction, for a listener is unable to pause, review what has gone before, or scan ahead to see what is coming.

3. Some of the usual rules of writing can be ignored in the interest of rapid comprehension. For example, you can write non-sentences and use abbreviations and contractions freely.

4. Within reason, what you write can be imprecise and incomplete—and even incorrect. These tactics are used to simplify the content of a slide, and to avoid excessive detail. Of course, to make sure that no confusion arises you must elaborate and explain the hidden or falsified features as you talk through the slide.

10.2. Designing the Talk

The first step in writing a talk is to analyse the audience. Decide what background material you can assume the listeners already know and what material you will have to review. If you misjudge the listeners' knowledge, they could find your talk incomprehensible at one extreme, or slow and boring at the other. If you are unsure of the audience, prepare extra slides that can be included or omitted depending on your impression as you go through the talk and on any questions received.

The title of your talk should not necessarily be the same as the one you would use for a paper, because your potential audience may be very different from that for a paper. To encourage non-specialists to attend the talk keep technical terms and jargon to a minimum. I once gave a talk titled "Exploiting Fast Matrix Multiplication within the Level 3 BLAS" in a context where non-experts in my area were among the potential audience. I later found out that several people did not attend because they had not heard of BLAS and thought they would not gain anything from the talk, whereas the talk was designed to be understandable to them. A better title would have been the more general "Exploiting Fast Matrix Multiplication in Matrix Computations".

A controversial title that you would be reluctant to use for a paper may be acceptable for a talk. It will help to attract an audience and you can qualify your bold claims in the lecture. Make sure, though, that the content lives up to the title.

It is advisable to begin with a slide containing your name and affiliation and the title of your talk. This information may not be clearly or correctly enunciated when you are introduced, and it does no harm to show it again. The title slide is an appropriate place to acknowledge co-authors and financial support.

Because of the fixed path that a listener takes through a talk, the structure of a talk is more rigid than that of a paper. Most successful talks follow the time-honoured format "Tell them what you are going to say, say it, then tell them what you said." Therefore, at the start of the talk it is usual to outline what you are going to say: summarize your objectives and methods, and (perhaps) state your conclusions. This is often done with the aid of an overview slide but it can also be done by speaking over the title

- • Introduction and Motivation
- ▷ Deriving Partitioned Algorithms
- • Block LU Factorization and Matrix Inversion
- • Exploiting Fast Matrix Multiplication

Figure 10.1. Contents slide: the triangle points to the next topic.

slide. The aim is to give the listeners a mental road-map of the talk. You also need to sprinkle signposts through the talk, so that the listeners know what is coming next and how far there is to go. This can be done orally (example: "Now, before presenting some numerical examples and my overall conclusions, I'll indicate how the result can be generalized to a wider class of problems"), or you can break the talk into sections, each with its own title slide. Another useful technique is to intersperse the talk with contents slides that are identical apart from a mark that highlights the topic to be discussed next; see Figure 10.1.[17]

Garver [102] recommends lightening a talk by building in multiple entry points, at any of which the listener can pick up the talk again after getting lost. An entry point might be a new topic, problem or method, or an application of an earlier result that does not require an understanding of the result's proof. Multiple exit points are also worth preparing if you are unsure about how the audience will react. They give you the option of omitting chunks of the talk without loss of continuity. A sure sign that you should exercise this option is if you see members of the audience looking at their watches, or, worse, tapping them to see if they have stopped!

An unusual practice worth considering is to give a printed handout to the audience. This might help the listeners to keep track of complicated definitions and results and save them taking notes, or it might give a list of references mentioned in the talk. A danger of this approach is that it may be seen as presuming the audience cannot take notes themselves and are interested enough in the work to want to take away a permanent record. Handouts can, alternatively, be made available for interested persons to pick up after the talk.

[17]To save space, all the boxes that surround the example slides in this chapter are made just tall enough to hold the slide's content.

10.3. Writing the Slides

Begin the talk by stating the problem, putting it into context, and motivating it. This initial scene-setting is particularly important since the audience may well contain people who are not experts in your area, or who are just beginning their research careers.

The most common mistake in writing a talk is to put too much on the individual slides. The maxim "less is more" is appropriate, because a busy, cluttered slide is hard for the audience to assimilate and may divert their attention from what you are saying. Since a slide is a visual aid, it should contain the core of what you want to say, but you can fill in the details and explanations as you talk through the slide. (If you merely read the slide, it could be argued that you might as well not be there!) There are various recommendations about how many lines of type a slide should contain: a maximum of 7–8 lines is recommended by Kenny [149] and a more liberal 8–10 lines by Freeman et al. [86]. These are laudable aims, but in mathematical talks speakers often use 20 or more lines, though not always to good effect.

A slide may be too long for two reasons: the content is too expansive and needs editing, or too many ideas are expressed. Try to limit each slide to one main idea or result. More than one may confuse the audience and weaken the impact of the points you try to make. A good habit is to put a title line at the top of each slide; if you find it hard to think of an appropriate title, the slide can probably be improved, perhaps by splitting it into two.

Don't present a detailed proof of a theorem, unless it is very short. It is far better to describe the ideas behind the proof and give just an outline. Most people go to a talk hoping to learn new ideas, and will read the paper if they want to see the details.

When a stream of development stretches over several slides, the audience might wish to refer back to an earlier slide from a later one. To prevent this you can replicate information (an important definition or lemma, say) from one slide to another. A related technique is to build up a slide gradually, by using overlays, or simply by making each slide in a series a superset of the previous one. (The latter effect can be achieved by covering the complete slide with a sheet of paper, and gradually revealing the contents; but be warned that many people find this peek-a-boo style irritating. I do not recommend this approach, but, if you must use it, cover the slide *before* it goes on the projector, not after.) Overlays are best handled by taping them together along one side, and flipping each one over in turn, since otherwise precise alignment is difficult.

If you think you will need to refer back to an earlier slide at some particular point, insert a duplicate slide. This avoids the need to search through the pile of used slides. It is worth finding out in advance whether two projectors will be available. If so, you will have less need to replicate material because you can display two slides at a time.

It is imperative to number your slides, so that you can keep them in order at all times. At the end of the talk the slides will inevitably be jumbled and numbers help you to find a particular slide for redisplaying in answer to questions. I put the number, the date of preparation, and a shortened form of the title of the talk, on the header line of each slide.

When you write a slide, aim for economy of words. Chop sentences mercilessly to leave the bare minimum that is readily comprehensible. Here are some illustrative examples.

Original:	It can be shown that $\partial\|Ax\| = \{A^T\mathrm{dual}(Ax)\}$.
Shorter:	Can show $\partial\|Ax\| = \{A^T\mathrm{dual}(Ax)\}$.
Even shorter:	$\partial\|Ax\| = \{A^T\mathrm{dual}(Ax)\}$.
Original:	The stepsize is stable if $p(z)$ is a Schur polynomial, that is, if all its roots are less than one in modulus.
Shorter:	Stepsize is stable if $p(z)$ is a Schur polynomial (all roots less than one in modulus).
Even shorter:	Stepsize stable if $p(z)$ is Schur ($\|\mathrm{roots}\| < 1$).
Original:	We can perform a similar analysis for the Lagrange interpolant.
Shorter:	Similar analysis works for the Lagrange interpolant.
Even shorter:	OK for Lagrange.

Lists are an effective way of presenting information concisely. Audiences like them, because there is something psychologically appealing about being given a list of things to note or remember. Consider using bulleted or numbered lists of five or fewer items, expressed with parallel constructions where appropriate (see §4.16).

If you cite your own work you can reduce your name to one letter, but the names of others should not be abbreviated. Thus, from one of my slides:

Original:	This question answered in nonsingular case by Higham and Knight (1991).
Shorter:	Answered for nonsingular A by H & Knight (1991).

Be sure to avoid spelling mistakes on the slides. They are much more noticeable and embarrassing in slides than in a paper.

Graphs and pictures help to break up the monotony of wordy slides. But beware of "slide shuffling", a common mistake where a long sequence of similar slides is displayed in rapid succession ("that was the plot for $\rho = 0.5$, now here is $\rho = 0.6$, ..., and here is $\rho = 0.7$,..."). An audience needs time to absorb each slide and it is difficult for anyone to remember more than the last two slides. Anything that might irritate or bore your audience is to be avoided.

At the end of the talk, conclusions and a summary are usually given. This is your last chance to impress your ideas upon the audience. If possible, try to summarize what you have said in a way that does not presuppose understanding of the talk. Even if your conclusions are largely negative, try to finish the talk on a positive note! The last slide is a good place to give your email address and a Web address where your relevant papers can be found.

Legibility of the Slides

It is of paramount importance that slides be legible when they are projected. Unfortunately, slides are frequently difficult to read for people sitting towards the rear of the room.

An important point to appreciate is that many overhead projector installations exhibit keystoning, whereby the image is wider at the top than at the bottom; this is avoided if the screen is tilted forward. Keystoning causes characters at the bottom of the screen to look much smaller than those at the top.

Take care not to write too close to the edges of a slide, as the areas near the edges may not be clearly visible when the slide is projected. And try to leave a sizeable gap at the bottom of the slide, as the bottom of the screen may be obstructed from the view of people at the back of the audience. I recommend experimenting with a projector to determine acceptable limits.

Various recommendations have been given in the literature on the minimum height of characters, but they necessarily involve assumptions about the magnification of the original image. My advice is to experiment with a projector in an empty room of a similar size to the one that will be used for the talk—you can quickly determine how large the characters need to be. I find that

characters this large (`\LARGE` in LaTeX at 10pt)

produce a slide readable from the back of most lecture rooms. For the comfort of the audience I prefer to use characters this large (\huge in LaTeX at 10pt).

For a given character size the font also affects readability [240, Chap. 2]. Some people prefer to use sans serif fonts for slides (fonts without small extensions at the ends of strokes):

sans serif font

It is not uncommon to see slides that have been photocopied from a paper without magnification. Invariably the slides are difficult or impossible to read. Avoid this sin; prepare concise, legible slides specially for the talk. The practice of photocopying from a paper is most common with tables. Better even than magnifying the original table is to rewrite it, pruning it to the information essential for that slide. For example, some entries in a table of numbers may be dispensable, and it may be possible to reduce the number of significant digits. More so than in a paper, a table of numerical results may be better displayed as a graph, which is easier to absorb as long as the labels are readable.

How Many Slides?

The number of slides needed depends on several factors, such as the length of the talk, the amount of mathematical detail to be presented, and the number of graphs and pictures. It also depends on the mode of presentation: a speaker who uses an overhead projector to display outline notes and who spends considerable time filling in the details on the blackboard will need far fewer slides than one who uses only the projector. In my experience, it is best to allow about two minutes per slide for "slide only" talks (the same recommendation is made in [86] and [102]). Allowing for variation in slide length, for a twenty minute talk you should aim for twelve slides or fewer, while for a fifty minute talk thirty slides is about the upper limit.

Handwritten or Typeset?

Which are better: handwritten or typeset slides? It is a matter of personal preference. The best handwritten slides are just as good as the best typeset ones, and the worst of both kinds are very poor. I produce all my slides in LaTeX, for several reasons. I can work from the LaTeX source for the corresponding paper, which saves a lot of typing; typeset slides can easily be adapted and reprinted for use in other talks; I find it easier to revise and correct typeset slides than handwritten ones, and this encourages me to produce better slides; and I have no worries about spilling liquid over my slides or losing them (they are never further away than an ftp call to my workstation). Advantages of handwritten slides are that they can be modified right up until the start of the talk without the need for access to a computer or printer, it is easy and inexpensive to use colours, and handwritten slides are more individual.

LaTeX 2_ε has a document class `slides` [172, §5.2] designed for making slides (it supersedes SLiTeX, a version of LaTeX 2.09). It automatically produces large characters in a specially designed sans serif font. It supports the production of coloured slides, via the `color` package, and overlays (so that one slide can be superimposed on another). Various other macro packages are available for use with LaTeX, including the `seminar` package by Timothy Van Zandt (available from CTAN—see §13.5). Note that typeset slides are easier to read if they are not right-justified and if hyphenation is turned off.

To break the monotony of black and white typeset slides you can underline headings and key phrases with a coloured pen. You can also leave gaps to be filled during the talk, later wiping the slides clean, ready for the next time you give the talk.

If you use typeset slides you may have the option of projecting directly from a computer through a projection device, as opposed to using transparencies on an overhead projector. One advantage of working from the computer is that you can change the slides right until the last minute. Another is that programs such as Microsoft PowerPoint permit sophisticated techniques to be used, such as animation and building up a slide layer by layer. There are several disadvantages:

- The quality of computer projection systems is variable. Unless a sharp, bright image is produced, it will be better to use transparencies.

- There is a tendency for authors to emphasize presentation at the expense of content when using the fancier presentation programs. One seminar organizer I know has banned the use of PowerPoint presentations.

- More things can (and do) go wrong so, as a golden rule, have printed slides ready just in case!

10.4. Example Slides

In this section I give some examples of how to improve slides. In each case two slides are shown: an original draft and a revised version. Keep in mind that the slides would be printed in much larger type, and magnified further by the projector. Therefore, what appear here to be only minor flaws will be much more noticeable in practice.

The slide in Figure 10.2 gives some numerical results. The second version improves on the first in several ways.

- The definition of the matrix has been changed from a precise but cryptic form to a descriptive one.
- The quantities ρ^N, ρ^R, and γ_2 would have been defined on an earlier slide, but the audience may need to be reminded of what they represent. Therefore, a description in words has been added.
- Since only the order of magnitude of the error is relevant, the numbers are quoted to just one significant figure. This reduces the visual complexity of the table. To reduce the complexity further, the column headed $\rho^C(\widehat{y})$ has been removed, and the minimum number of rules has been used in the table.

In Figure 10.3, the original slide contains a theorem as it would be stated in a paper. The revised slide omits some of the conditions, which can be mentioned briefly as the slide is explained, and reminds the viewer of the definition of $\kappa_2(A)$. The simplified slide is much easier to assimilate.

The revised slide in Figure 10.5 omits some unnecessary words and symbols from the slide in Figure 10.4. The slide in Figure 10.6 has far too many words. Ruthless pruning produces the much more acceptable version in Figure 10.7; as usual, the information removed from the slide would be conveyed when giving the talk. Note that the parentheses in the expressions for β_{n+1} and π_{n+1} were unnecessary and so have been omitted, and the final error bound is written in a more readable way. Another possible improvement is to remove the definitions of α_0, β_0 and π_0, which may be irrelevant for this slide.

10.5. Further Reading

See §11.3 in the next chapter.

Example

The unit roundoff is $u \approx 1.1921 \times 10^{-7}$.

$$A = \text{randsvd}([10,16], 1e2)$$

$$\kappa_2(A) = 1.00e2,\ \text{cond}_2(A) = 8.63e1$$
$$\text{cond}_2(A,x) = 1.57e2$$

	$\rho^N(\widehat{y})$	$\rho^R(\widehat{y})$	$\rho^C(\widehat{y})$	$\gamma_2(\widehat{y})$
Q method	1.83e−8	9.88e−9	1.42e−7	2.01e−6
SNE	5.11e−7	2.79e−7	4.40e−6	4.97e−6
	1.52e−8	6.45e−9	9.64e−8	1.99e−6

Example

Unit roundoff $u \approx 1.2 \times 10^{-7}$.

A 10×16, random

$\kappa_2(A) = 1e2, \quad \text{cond}_2(A) = 9e1$

	ρ^N	ρ^R	γ_2
Q method	2e−8	1e−8	2e−6
SNE	5e−7	3e−7	5e−6
	2e−8	6e−9	2e−6

ρ^N = normwise backward error
ρ^R = row-wise backward error
γ_2 = forward error

Figure 10.2.

PERTURBATION BOUNDS

Traditional Result

Theorem Let $A \in \mathbb{R}^{m\times n}$ and $0 \neq b \in \mathbb{R}^m$. Suppose rank$(A) = m \leq n$ and that $\Delta A \in \mathbb{R}^{m\times n}$ and $\Delta b \in \mathbb{R}^m$ satisfy

$$\epsilon = \max\{\|\Delta A\|_2/\|A\|_2, \|\Delta b\|_2/\|b\|_2\} < \sigma_m(A).$$

If x and $\widehat{x}$ are the minimum norm solutions to $Ax = b$ and $(A + \Delta A)\widehat{x} = b + \Delta b$ respectively, then

$$\frac{\|\widehat{x} - x\|_2}{\|x\|_2} \leq \min\{3, n - m + 2\}\kappa_2(A)\epsilon + O(\epsilon^2).$$

PERTURBATION BOUNDS

$A \in \mathbb{R}^{m\times n}$, $m \leq n$.

$$\begin{aligned} x = &\text{ min norm sol'n to } Ax = b, \\ \widehat{x} = &\text{ min norm sol'n to } (A + \Delta A)\widehat{x} = (b + \Delta b). \end{aligned}$$

Traditional Bound

$$\frac{\|\widehat{x} - x\|_2}{\|x\|_2} \leq c_{m,n}\, \kappa_2(A)\epsilon + O(\epsilon^2),$$

where

$$\epsilon = \max\left\{\frac{\|\Delta A\|_2}{\|A\|_2}, \frac{\|\Delta b\|_2}{\|b\|_2}\right\},$$

$$\kappa_2(A) = \|A^+\|_2\|A\|_2.$$

Figure 10.3.

MACHIN'S ARCTAN FORMULA

John Machin (1680–1752) used Gregory's arctan series to obtain a more rapidly convergent series for computing π. He argued as follows.

If $\tan\theta = 1/5$, then

$$\tan 2\theta = \frac{2\tan\theta}{1-\tan^2\theta} = \frac{5}{12}$$

and

$$\tan 4\theta = \frac{2\tan 2\theta}{1-\tan^2 2\theta} = \frac{120}{119}.$$

Since $\tan 4\theta = 120/119$ and $\tan(\pi/4) = 1$, this suggests we evaluate

$$\tan(4\theta - \pi/4) = \frac{\tan 4\theta - 1}{1+\tan 4\theta} = \frac{1}{239}.$$

This gives

$$\begin{aligned}\arctan\frac{1}{239} &= 4\theta - \frac{\pi}{4} \\ &= 4\arctan\frac{1}{5} - \frac{\pi}{4},\end{aligned}$$

or

$$\pi = 16\arctan\frac{1}{5} - 4\arctan\frac{1}{239}.$$

In 1706 Machin used this formula to compute 100 digits of π.

Figure 10.4.

MACHIN'S ARCTAN FORMULA

John Machin (1680–1752): if $\tan\theta = 1/5$, then

$$\tan 2\theta = \frac{2\tan\theta}{1-\tan^2\theta} = \frac{5}{12}.$$

Similarly, $\tan 4\theta = 120/119$. Then

$$\tan(4\theta - \pi/4) = \frac{\tan 4\theta - 1}{1+\tan 4\theta} = \frac{1}{239}$$

$$\begin{aligned}\Rightarrow \quad \arctan\frac{1}{239} &= 4\theta - \frac{\pi}{4}\\ &= 4\ \arctan\frac{1}{5} - \frac{\pi}{4}.\end{aligned}$$

In 1706 Machin used Gregory's arctan series to compute 100 digits of π.

Figure 10.5.

Fast Methods for Computing π

Up to the 1970s all calculations of π were done using linearly convergent methods based on series, products, or continued fractions. For these, doubling the number of digits in the result roughly doubles the work.

Brent and Salamin independently discovered a quadratically convergent iteration for π in 1976. For such an iteration, the errors satisfy $e_{k+1} = O(e_k^2)$, so each iteration step approximately doubles the number of correct digits.

Borwein and Borwein (1984) discovered another quadratically convergent iteration, and later found even higher order iterations. Their 1984 iteration is

$$\alpha_0 = \sqrt{2}, \quad \beta_0 = 0, \quad \pi_0 = 2 + \sqrt{2},$$
$$\alpha_{n+1} = \frac{1}{2}(\alpha_n^{1/2} + \alpha_n^{-1/2}),$$
$$\beta_{n+1} = \alpha_n^{1/2}\left(\frac{\beta_n + 1}{\beta_n + \alpha_n}\right),$$
$$\pi_{n+1} = \pi_n \beta_{n+1}\left(\frac{1 + \alpha_{n+1}}{1 + \beta_{n+1}}\right),$$

for which $|\pi_n - \pi| \le \frac{1}{10^{2^n}}$.

Figure 10.6.

Fast Methods for π

Up to 1970s: linearly convergent series, products, or continued fractions.

Brent and Salamin (independently) 1976: quadratically convergent iteration for π. Errors satisfy $e_{k+1} = O(e_k^2)$.

Borwein and Borwein (1984, 1987): another quadratically convergent iteration and higher order ones. Quadratic iteration: $\alpha_0 = \sqrt{2}$, $\beta_0 = 0$, $\pi_0 = 2 + \sqrt{2}$,

$$\alpha_{n+1} = \tfrac{1}{2}(\alpha_n^{1/2} + \alpha_n^{-1/2}),$$

$$\beta_{n+1} = \alpha_n^{1/2} \frac{\beta_n + 1}{\beta_n + \alpha_n}, \quad \pi_{n+1} = \pi_n \beta_{n+1} \frac{1 + \alpha_{n+1}}{1 + \beta_{n+1}},$$

for which $|\pi_n - \pi| \le 10^{-2^n}$.

Figure 10.7.