

**Good ideas do not convince most people because they are good:
You have to convince them!**

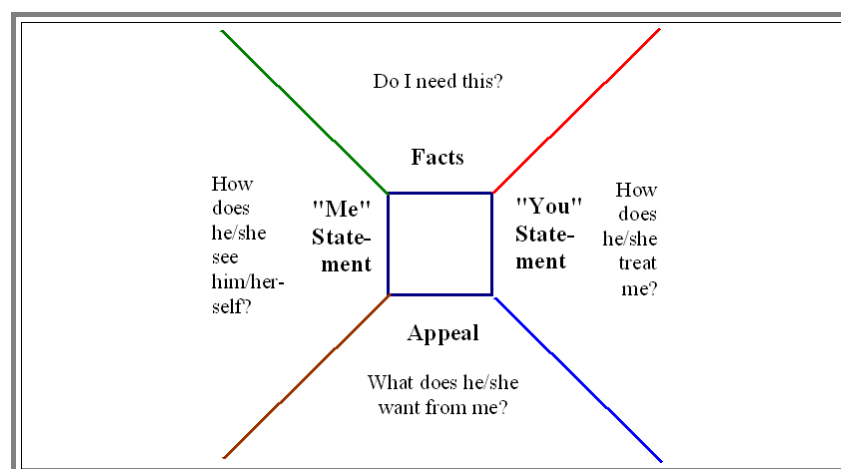
Advanced Oral Presentations


1. Communication



Some general Points




- ▶ The success of a presentation- however measured - will always depend on two factors:
 - **The factual content**
(here we discuss only scientific, not political presentations, so we assume there is content)
 - **The packaging**
- ▶ Both factors are equally important in a first approximation! Remember: **Success = Content · Acceptance** in lay terms, and **Success = (a₁ · Content)ⁿ · (a₂ · Acceptance)^m**, a_i, m, n = cost, if you want to be more sophisticated and put weight factors on the 2 ingredients, but for acceptance = 0, success will be = 0, anyway.
- ▶ Now, as far as content goes, try to remember presentations that you heard and found good. Most likely two conditions were met
 - You understood what was being said, at least, even if getting lost in details, you always could follow the **red line** of the presentation.
 - You were not put out by the packaging - you "liked" the speaker. The way he/she presented stuff kept you willing to process what you heard.
- ▶ The question is, how do you make sure that your audience feels that way about your presentation?
 - If you are not a fascinating person by definition (e.g. a Noble Prize Winner), you have to appeal to your audience on a factual and emotional level.
 - Since scientific presentation are supposed to be unemotional in the conventional sense of the word, your emotional impact must come from the way you speak, you move, you look at your audience, you formulate your sentences and so on.

Four Aspects of Communications










-  The recipient of your presentation has a completely free choice of what aspects he/she emphasizes for himself/herself. If there are many recipients, changes are that they will walk away and they all have heard quite different things if you ask them about the presentation a few days later. The message received by each individual differs from that of his neighbour, and **all** reception may differ from what was sent or from **what you think**, you sent.

 -  There is nothing you can do about this - except to make sure that on top of varying memories they all (or at least most of them) have the same recollection of just a few **essentials**.
-  And if you rack your own memory of some presentation you heard in the past, your recollection will always be along two totally separate lines

 -  You may remember something about the topic:
 ("...it had something to do with Si chips...")
 -  You remember something about the presentation:
 ("...he made a lot of jokes...", "...he was barely understandable..", "... he forgot to remove his bicycle clips from his pants...", "...I forgot what it was all about, but it was very interesting", "... was that the talk where everybody fell asleep?").
-  Try it! If you can remember any presentation without remembering something on this "emotional" level, you are actually dead and were replaced by an alien robot!

Noise in Communication

-  As you (should) know from communication theory, any communication channel may be disturbed by noise or other aberrations. Now, you are sending on **two** channels, the factual one and the emotional one. And even if the factual channel is noise-free, noise on the emotional channel influences the reception on the factual channel - there is heavy cross-talk!
-  Lets look at some of the reasons for noise

 -  **Incongruent signals:**
 Factual and emotional (or spoken and unspoken) messages differ. A trivial example: If you discuss equation X, but point at equation Y, your audience gets confused.
 -  **Unfavourable relations to audience**
 Being factually correct may be emotionally wrong. Saying repeatedly "...as you all should know from High school..." may be factually correct, but the people who forgot, will hate you and won't develop a positive attitude towards your message.
 -  **Being hard to follow**
 This can happen in quite different ways. If you say "...as is immediately apparent, the solution to this (incredibly long and complex) differential equation is $c = v \cdot \lambda$.. " you loose your audience (it is either insulted or thinks about why something is immediately apparent that is not), but you also loose it if you start solving your equation for a long time (the audience meanwhile forgets what the solution is good for).
 -  **Biased Recipients**
 They shouldn't exist in science, but then, we are all human. If you to explain to Prof. X and his crew, why their pet theory is all wrong, your audience will be biased and receive what you say heavily filtered. The same thing happens, just with signs reversed, if your stuff supports his pet theory. Students have accused their Professor of being against alternative energies, because he pointed out that there are only so much kWh that you may get from any solar cell in this solar system. The message received was completely different from what was sent because a bias developed early in the presentation.
 -  **Blocking**
 In the extreme form of the above issue, the recipient will simply no longer listen or turn everything around.

2. Body Language

Noise Sources on the Emotional Channel

There are a few "classical" sources of noise on the emotional channel, that may heavily interfere with the signals on the factual channel:

- **No eye contact.**
If you talk to the blackboard, to the overhead projector or to someone in outer space, you are going to loose your audience. **That is not easy to avoid.** Sometimes it is helpful if you pick a few persons (not too close up) in the audience, to whom you talk keeping eye contact (for only a few seconds each!).
- **Nervously running back and forth**
- **Standing stiffly in one place all the time**
- **Lots of gesticulation**
- **No gesticulation**
If you secretly hope that your hands will disappear because you don't know what to do with them, your audience will notice (**very difficult problem!**). In scientific talks there are simple tricks: Writing on foils or on the blackboard, keeping a pointer in your hand (but then you must use it sensibly).
- **Playing with the pointer**
- **Fumbling around a lot with your notes**
- **Worst of all: reading everything from your notes**
(while looking at your notes)

How to Appear Secure

It doesn't matter if you feel secure and confident, it matters if your audience gets that impression. However, it is a lot easier to convey that impression if you actually **are** secure and confident. But there are tricks:

- **Stand securely,**
legs slightly spread, erect and face your audience. (if you're a female, don't wear high heels if it is not a fashion event)
- **Control your gesture,**
(but forget that if you are a beginner). Still, the idea is to move your arms only above the belt line and outside of the chest area.
- **Be loud, be slow and make pauses**
A loud voice (not screaming) is a signal of security. Machine gun speech patterns are only in character if it is one of your trade marks (beginners have no trade marks).
- **Controlled position changes**
Walk calmly to the overhead projector, blackboard etc. If you are extremely controlled, make a little stop on your way to wherever and continue your presentation with a few remarks. That requires that you start to walk **before** the issue that demands the walk comes up!
- **Calmly face your "contact persons" eye to eye - but no longer than 3 seconds.** Have at least three contact persons or segment of the audience between which you change your eye contact.

But don't forget: If you actually try to remember and do all this on your **first few** presentations, you will definitely forget what you wanted to talk about (all this needs practice). **If you neglect the signals on the factual channel, zero negative interference on the emotional channel cannot have any positive impact!**

[Next page](#)


Several pages to the same subject will follow





[Back to the "Guided Tour" home page](#)




Good ideas do not convince most people because they are good: You have to convince them!



1. Communication

Some general Points

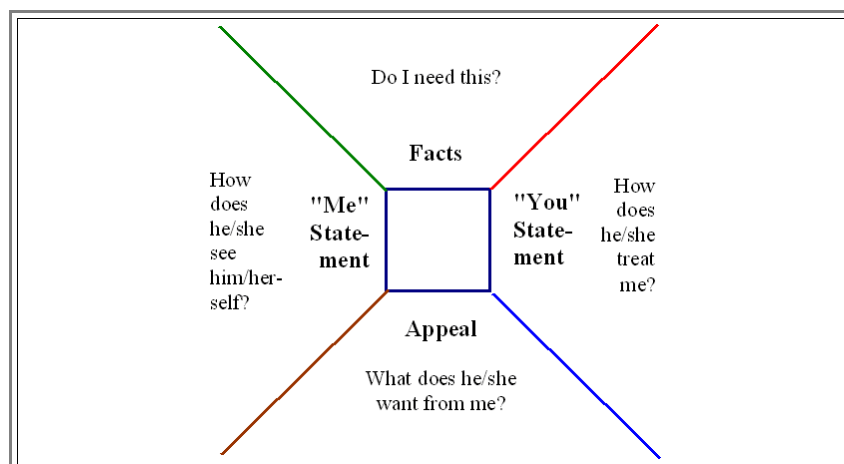
-  The success of a presentation- however measured - will always depend on two factors:


 -  *The factual content*
 (here we discuss only scientific, not political or sociological presentations, so we assume there *is* content)
 -  *The packaging*
-  Both factors are equally important in a first approximation! Remember: $\text{Success} = \text{Content} \cdot \text{Acceptance}$ and that means that for acceptance = 0, success will be = 0, too - even if you talk about the work that will get you the Noble prize.
-  Now, as far as content goes, try to recall some presentations that you heard and found good. Most likely two conditions were met:



 -  You *understood* what was being said. Even if you got lost on occasion, you always could follow the *red line* of the presentation.
 -  You were not put off by the *packaging* - you sort of *liked* the speaker. The way he/she presented the stuff kept you willing to continue listening.
-  The question is, how do *you* make sure that your audience feels that way about *your* presentation?




 -  If you are not a fascinating person by definition (e.g. a Noble Prize Winner, or [Brittney Spears teaching semiconductor physics](#)), you have to appeal to your audience on a factual *and* emotional level.
 -  Since scientific presentation are supposed to be *unemotional* in the conventional sense of the word, your emotional impact must come from the way you speak, you move, you look at your audience, you formulate your sentences, and so on.

Four Aspects of Communications






-  The recipient of your presentation has a completely free choice of what aspects he/she emphasizes for himself/herself. If there are many recipients, changes are that they will walk away and they all have heard quite different things if you ask them about the presentation a few days later. The message received by each individual differs from that of his neighbour, and *all* receptions may differ from what was sent, or from *what you think* you sent.






 -  There is nothing you can do about this - except to make sure that on top of varying memories they all (or at least most of them) have the same recollection of just a few *essentials*.
-  And if you rack your own memory of some presentation you heard in the past, your recollection will always be along two totally separate lines

 -  You may remember something about the *topic*:
 ("...it had something to do with Si chips...")
 -  You remember something about the *presentation*:
 ("...he made a lot of jokes...", "...he was barely understandable..", "... he forgot to remove his bicycle clips from his pants...", "...I forgot what it was all about, but it was very interesting", "... was that the talk where everybody fell asleep?").
-  Try it! If you can remember any presentation without remembering something on this "emotional" level, you are actually dead and were replaced by an alien robot!

Noise in Communication

-  As you (should) know from communication theory, any communication channel may be disturbed by noise or other aberrations. Now, you are sending on *two* channels simultaneously - the factual one *and* the emotional one.

 -  And face it: Even if the *factual* channel is noise-free, noise on the emotional channel influences the reception on the factual channel - there is heavy cross-talk!
-  Lets look at some of the reasons for *noise*:

 -  *Incongruent signals*:
 Factual and emotional (or spoken and unspoken) messages differ. A trivial example: If you discuss equation X, but point at equation Y, your audience gets confused.
 -  *Unfavourable relations to audience*
 Being factually correct may be emotionally wrong. Saying repeatedly "...as you all should know from High school..." may be factually correct, but the people will start to hate you after the third time and won't develop a positive attitude towards your message.
 -  *Being hard to follow*
 This can happen in quite different ways. If you say "as is immediately apparent, the solution to this (incredibly long and complex) differential equation is $c = v \cdot \lambda$..", you loose your audience (it is either insulted or thinks about why something is immediately apparent that is not), but you also loose it if you start solving your equation for a long time (the audience meanwhile forgets what the solution is good for).
 -  *Biased Recipients*
 They shouldn't exist in science, but then, most of us are human. If you to explain to Prof. X and his crew, why their pet theory is all wrong, your audience will be biased and receive what you say heavily filtered. The same thing happens, just with signs reversed, if your stuff supports his pet theory. Students have accused their Professor (me) of being against alternative energies, because he pointed out that there are only so much kWh that you may get from any solar cell in this solar system. The message received was completely different from what was sent because a bias developed early in the presentation.
 -  *Blocking*
 In the extreme form of the above issue, the recipient will simply no longer listen or turn everything around.

2. Body Language

Noise Sources on the Emotional Channel

There are a few "classical" sources of noise on the emotional channel, that may heavily interfere with the signals on the factual channel:

- *No eye contact.*
If you talk to the blackboard, to the overhead projector or to someone in outer space, you are going to lose your audience. *That is not easy to avoid.* Sometimes it is helpful if you pick a few *contact persons* (not too close up) in the audience, to whom you talk by keeping eye contact (for only a few seconds each!).
- *Nervously running back and forth all the time.*
- *Standing stiffly in one place all the time.*
- *Lots of gesticulation.*
- *No gesticulation.*
If you secretly hope that your hands will disappear because you don't know what to do with them, your audience will notice (*very difficult problem!*). In scientific talks there are simple tricks: Writing on foils or on the blackboard, keeping a pointer in your hand (but then you must use it sensibly).
- *Playing with the pointer.*
- *Fumbling around a lot with your notes.*
- *Worst of all: reading everything from your notes*
(while looking at your notes)

How to Appear Secure

It doesn't matter if you *feel* secure and confident, it only matters that your audience gets that impression. However, it is a lot easier to convey that impression if you actually *are* secure and confident. But there are tricks:

- *Stand securely.*
Legs slightly spread, erect and face your audience. (if you're a female, don't wear high heels if it is not a fashion event).
 - *Control your gestures.*
(but forget that if you are a beginner). Still, the idea is to move your arms only above the belt line and outside of the chest area.
 - *Be loud, be slow and make pauses.*
A loud voice (not screaming) is a signal of security. Machine gun speech patterns are only in character if it is one of your trade marks (beginners have no trade marks).
 - *Controlled position changes.*
Walk calmly to the overhead projector, blackboard etc. If you are extremely controlled, make a little stop on your way to wherever, and continue your presentation with a few remarks. That requires that you start to walk *before* the issue that demands the walk comes up!
 - *Calmly face your "contact persons" eye to eye - but no longer than 3 seconds. Have at least three contact persons or segments of the audience between which you change your eye contact.*
- But don't forget: If you actually try to remember and do all this on your *first few* presentations, you will definitely forget what you wanted to talk about (all this needs practice).
- *If you neglect the signals on the factual channel, zero negative interference on the emotional channel cannot have any positive impact!*

[Next page](#)

3. How to Prepare a Presentation

The Time Needed for a Good Presentation

Advanced

It is conventional wisdom, that **95%** of the work needed for a good presentation is done *before* you face your audience.

- A rule of thumb for *experienced* speakers, well versed in their field, is:

30 minutes preparation for 1 minute presentation!

Included in that time are:

- Basic decisions.**
Analysis of the likely structure and expectations of the target group (is your target group the Prof. and his assistants or your fellow students?), first decisions about the main goals of the presentation (what are the main messages going to be), decision on media use (blackboard, overhead projector, flip chart, laptop and beamer, small experiments, objects to be shown around, ...).
- Conception.**
Collecting materials, basic structure of the presentation, how to visualize certain points ...
- Production.**
Making your viewgraphs, notes ...
- Control.**
Trying out the viewgraphs, talking to the mirror to find out how long it takes ...
- Dress rehearsal.**
Actually giving the presentation to a few good friends in a suitable environment (not at your apartment with the TV in the background and everybody drinking beer).

Intentions and Goals

This is the important part. Don't forget: All the good advice about the emotional stuff will do no good if the factual content of your presentation is no good! There is one golden rule:

- If you don't know exactly where you want to go to, you shouldn't be surprised if you don't arrive there.**

The first question you have to ask yourself in preparing a presentation is:

- What do I want to get across? Which messages do I want to implant firmly into the brains of my audience?

And do not forget: A presentation is *not* the same thing as a lecture. Your audience is not going to really learn something from you - people never learn anything from a lecture either, they learn it by mulling the issues over themselves. ilivate, in exercise classes, in discussions, in preparing for an exam - but not during the lecture.

- The audience listening to a presentation will, in general, *not* do this!
- The sad point about this is that there is very little that you can get across that will "stick" for some time. So keep you message to a *few bare essentials*..
- There are very experienced speakers out there who claim that the number of messages that can be firmly delivered in *one* presentation is exaxctly 1!

Tuning in to the Receiver

- ▶ Know your audience! Ask yourself: What do I have to take into account with *this group* of listeners, to get my messages across? Ask yourself relative to your target group (which does not have to be all of the audience, but maybe only some key persons):
 - How does my audience (probably) see the topic?
 - What is their basic attitude? Do they love the issue, but hate to be here at this time, ...?
 - What does the audience expect from this presentation?
 - What is their knowledge background (Even Professors hate sentences like: "I would be insulting your intelligence by explaining the mass action law in this context; it is, after all, basic high school stuff...")

4. The Introduction

General

- ▶ As in aviation, there are two critical phases in a presentation: *Starting* and *landing*.
 - Your start - the *introduction* - determines if you keep your audience interested. The audience wants to recognize early on if this is interesting enough to concentrate on.
 - Ideally your introduction should do two things:
 - Make the listener curious.
 - Signal your competence.
- ▶ What can make your *start* interesting, if not outright thrilling? Easy: *Don't do what your audience expects!*.
 - Ask a surprising (rhetorical) question.
 - Find a connection to a recent big event.
 - Start with a good quote.
 - Give a surprising piece of information.
 - Start with a provocation.

Themes and Theses

- ▶ A good presentation has a *theme* which can be developed by postulating *theses*.
- ▶ A good theme should be short and with no additional explanation. We are *not* discussing the *title* of the presentation, but the *key sentences* at the beginning!
 - A *good* theme sentence:
I am going to demonstrate that fuel cells are the *only* option to preserve *your* present mobility in the near future.
 - A *bad* theme sentence:
As you may know, fuel cells, provided we will be able to make them cheaply enough, will be important in solving some of the problems of transportation sometime in the future, always assuming that nuclear fusion, as we have reason to fear, will not make it, and, of course, if we can get a hydrogen economy going in time.
- ▶ Always ask yourself if your theme sentence(s) can be formulated even more succinctly.

Theses

- ▶ Theses are the central points of any presentation where you are trying to "*sell*" something, or, more general, to *induce some action* of your audience after the speech.
 - The action may be that they buy your product, give you a job, or elect you to the Bundestag.
 - However, in *scientific presentations*, you may only explain something that your audience didn't know and you do not expect actions. So this part about the theses must be taken with a grain of salt!
- ▶ If you do use theses, start with the conclusions! There are several advantages:
 - If you show your intentions right away, you appear more trustworthy.
 - If the audience knows where you are heading to, they have an easier time following you.
- ▶ A thesis expressing *your* opinion, could start as follows:
 - I am of the opinion that..
 - I think that...

- I am convinced that...

- ▣ A thesis sentence *never* contains

- Explanations.
- Examples.
- Justifications.

What Belongs to an Introduction?

- ▣ *Welcoming* the audience, and, depending on the occasion, *introducing* yourself.

- ▣ A *starting part* with your surprises, provocations, quotes etc.

- ▣ A *theme* sentence.

- ▣ The *theses*.

- ▣ The *background information* (can be very short); e.g.

- Relevance of your theme.
- History (in Germany always refer to the old Greeks).
- Personal relation to the theme.

- ▣ An *orientation* (this is absolutely *de rigueur* in a scientific talk!). It should contain most, if not all of the following:

- Contents and structure.
- Time plane.
- What kind of documents are handed out to the participants (You must tell them if they should take notes, or if they do not have to bother). *Never* give your documents out before you are finished!

- ▣ The introduction should be within about **15%** of the total time allowed for the presentation!

[Previous page](#)

[Next page](#)

5. Main Part

Structuring a Theme

Advanced

- For general presentations (especially of the selling type), the basic approach for a structure is the "Now - Then" comparison.
 - Start with "*Now*". What is the situation now, what is good about it, what is not so good, what are the problems.
 - Continue with describing "*Then*". What could be, what would be optimal, what kind of advantages are expected (especially for your audience).
 - Now present your *solutions*, or how we get from now to then. Show alternatives, give reasons for your choice between alternatives, discuss pros and cons (be especially open about the cons), end with the most important pro argument.
- This recipe can also be used (at least in parts) for scientific presentations where you present your *research results*. For scientific presentations where you just *explain* existing knowledge to your not yet enlightened audience, this scenario may not be very helpful.

Being Convincing

- Even in science, you have to be convincing if you want to be believed, or even more important, if you want to be remembered.
- Now facts are facts, but how convincing they are depends on how you present them. Be concrete, even dramatic (but don't overdo it with a scientific audience). Examples:
 - True, but not very convincing:*
Quartz oscillators are very important for many electronic products.
 - True, and attention getting:*
If the production of quartz oscillators would stop for some reason, the computer and communication industry would completely collapse within weeks.
 - True, and (overly) dramatic:*
If quartz oscillators for some unknown reason were to suddenly stop functioning, millions of people would die within hours!
- Show hard facts. Answer obvious or suspected questions along the following lines *before* they are asked:
 - Who says so?
 - What right do you have..?
 - Who proves that this is true?
- We believe what we *see* much more than what we *hear*!
 - Illustrate your points; visualize!

6. The Ending

The Importance of a Good Ending

- Ende gut, alles gut! (German saying: "All is well that ends well"). There are reasons for that:
 - The attention level always goes up before the ending (if only because people start to wake up).
 - You can once more formulate your message and hammer it home.
 - You can once more make clear (directly or indirectly) what the audience is supposed to do (give you a good grade, for example).
(..must say that that the topic was a big challenge, but I have learned a lot working on it and enjoyed it very much. I hope that in the next years students will also get an opportunity to go through this great experience....).
 - With a good ending you create (hopefully, once more) the impression that you are an expert in your field, a professional in whatever you do, and on top of it a good speaker!
- Don't just fade away because the time is up, or you run out of things to say, *plan your ending!*

How to Make a Good Ending

- There are some time-proved endings (not all applicable to scientific talks):
 - *Announcement:*
Let me finally summarize: ..
 - *Conclusion:*
In conclusion, the main problem is..
The proposal for further action thus is...
To summarize, the following results were obtained..
 - *Moving to the discussion:*
I expect questions to this point and am perfectly willing to answer...
- You should allow about **10 %** of your time for the end part of your presentation.

[Previous page](#)

[Next page](#)

7. Manuscript

The Manuscript for the Presentation

Advanced

- There is a simple rule: *There is no manuscript for the presentation!*. *Never ever* read from a paper while giving a scientific presentation! However:
 - It is useful to write down some of what you are going to say for yourself!
 - It is admissible to write down the important opening or closing sentences and to look at them when they come up.
 - It is recommended to write down a "skeleton" of your presentation to which you can refer. This - if you want to do it "professionally" - could be in the form of **DIN A 6** cards (one for very topic) which you shuffle as the presentation rambles on; watch showmaster or politicians - they all use that system.
- If you do that, keep in mind:
 - Keep your manuscript (with the skeleton) or your cards *horizontal* - the area seen by the audience is smaller this way, it is less obtrusive.
 - The hand that is holding the cards stays quiet. Gesticulations are for the other hand!
- In scientific talks there is an alternative: Write your skeleton on viewgraphs ("Skeleton foils")
 - Mark on those viewgraphs places where you are going to show another viewgraph, write something on the blackboard, etc.
 - But don't forget to go back to the proper place on your skeleton viewgraph.
 - And do not mix up your viewgraphs!
- An [example](#) for a "skeleton foil" is given in the link. The cartouches mark an action. "Folie" means that a different viewgraph will be shown at this point.

The Paper to the Presentation

- The paper to the presentation (the hand out after the presentation) is *not* the written version of the presentation!
- Of course it contains everything you said and showed, but it may contain a lot more. In fact it *has* to contain more:
 - It must have the *names of the authors* on it, a date and possibly some other comments to the w's (why, when, where, who, what for, with what, for whom, ...)
 - It must have the *Figures*, and if they are not your own, their origin.
 - It may contain any amount of *formulas* (which your presentation should not!).
 - It must have *references* to other papers and books, and a list of references at the end.
 - It should have an "*Acknowledgment*" at the end if you have reason to mention any help you got from others in writing the paper.
(... I'm especially indebted to Ms. Anderson who regularly provided delightful diversions which helped to clear my mind for the demanding task of writing this paper, and to the Holsten brewery whose products induced the necessary peace of mind.)
- In other words, it should be written in the time-honoured style of any scientific paper.

8. Visualization

General Remarks

One picture says more than a thousand words - you have heard that before (have you ever seen it?).

It is trivial, but it is true! If you use pictures, graphics, photographs, even viewgraphs with plain text on it, the effects are:

- Your information will stick better. Information processing in the brain is more efficient through the eyes than through the ears.
- You will appear more convincing and more trustworthy (ever noticed that the military, when they show a big victory, now always presents a satellite photography or something else where you actually cannot see anything clearly, but it is still very convincing!).

Everything you offer your audience to look at (including yourself) is a visual aid to your presentation. But not all possible visualizations are good visualization!

Keep in mind: Visualization are used to support your factual content, especially your key points. When conceiving of a possible visual aid, ask yourself the following questions:

- Is the visualization helpful at guiding the audience to where you want them to go?
- Is the visualization helping the listeners to understand complex relations that you are trying to point out?
- Is the visualization helpful to regain the attention of your audience at some specific point in your presentation?

From the Idea to the Visualization

First you have to select the information that is especially important to *the audience*. That may be different to what was especially important to *you*!

- If, for example, you spent a lot of time understanding how a particular equation or an integral was solved, the way to the solution may have been important to you. But it is probably not important for your audience! All that matters may be the boundary and starting conditions and the result. In this case do *not* visualize the math! Do *not* use a viewgraph with lots of formulas on it!
- Generally speaking: Mostly the details don't matter, but the consequences from the details. If the only information you actually use from a detailed table of something is the fact that item "X" accounts for about $\frac{1}{3}$ of whatever it is, do not show the table!
- But now let's assume you do have very interesting information and that you definitely are going to visualize it. Then you must decide what will be the best way. An example: Let's assume you have the following table giving the facts.

Year	1973	77	81	84	87	90	95	98	01	05	09	13	17
Price for 1 Mbit of DRAM memory (DM)	150.000	10.000	800	240	60	10	1	0,26	0,11	0,05	0,014	0,008	0,003

- Should you show it? The answer is an emphatic *no*! This can be [visualized much better](#) (with, however, a lot of additional work!).

The less abstract you make your data, the more you appeal directly to the emotional part of your audience - and that means that the information does not have to be processed to leave a clear imprint in the memory of the brain.

[Previous page](#)

[Next page](#)

8.1. Tips for Visualization

Don't Know How

Advanced

- ▶ You cannot come up with an idea for visualizing a certain subject because it is very general or very abstract?
 - Consider illustrating the *consequences* of that topic for a specific example.
 - Take *one* part for the whole and illustrate what happens then.
 - *Forget it.* Illustrations that are obviously only included for illustrations sake, but do not really help to make your point, are counterproductive!

Checklist for Visualisation

- ▶ Go through the following points for every illustration you consider:
 - Which idea should be communicated?
 - What kind of format is optimal (photography, graphic pictures, diagrams, tables...)?
 - Is the illustration supporting the idea or is it included because you have it, or it's such a neat picture?
 - Is the illustration stimulating? Intellectually or emotionally?
 - Does the illustration allow you some leeway for explaining? A totally self-explaining illustration is a bad illustration.
 - Is the illustration with your explanations clear and understandable?
(If you have to say: "...and also ignore the table in the lower hand corner and mentally substitute magnetic field strength **B** for wherever you see electrical fields strength **E**...", it is a lousy illustration!)
 - Does the format match the purpose?
A three-dimensional perspective drawing that clearly took hours to make is not a good match for illustrating simple things were one dimension would have been all that is needed.
 - Is the illustration within your general level of sophistication?
A black-and-white table quickly copied on a foil will look totally out of place if everything else is colourful and very sophisticated. It also works the other way around. Try to keep one (your!) standard throughout your presentation.

Texts and Tables

- ▶ There are a few very important points about how to write on viewgraphs!
- ▶ **Readability.** Whatever is written, must be readable from *all places* in the audience!
 - *Never* use typewriter fonts and size, i.e. font **10** or **12**!
 - Minimum letter size on a viewgraph is ca. **5 mm**; this corresponds to a font size of *at least 14 bold*, better **18**.
 - There are reasons for
 - Black on white.
 - Make sure to provide enough contrast between the letters and the background.
 - If you *have to* go to a smaller font because otherwise it won't fit on the foil, you have *too much. Never*, really *never*, put more on a foil as will fit with font **16**, at the very minimum font **14**.

Clarity.

- If the audience has to exert it's mental capability to try to understand what it sees on your illustration, they will not listen to what you say!
- The biggest enemies to clarity are *volume* and *precision*! Complete and precise information (with all the little disclaimers, validity ranges, boundary conditions and exceptions to the general rule) belong in the handout, not on the viewgraphs!

- Guide the attention to the core information! Generally, the audience should be able to grasp the contents of a viewgraph within **30** seconds. There may be exceptions if you work with the illustration, e.g. by overlaying it with other viewgraphs.
- 'Stay within one format! Use the same colours or symbols throughout your presentations for the same effects.

Attractiveness

- The choice of fonts and colours, of line sizes, frames etc. determines to a large extent if your viewgraph looks attractive.
- Of course, beauty rests in the eye of the beholder, but there is a general consensus.
- Use colours *sparingly* and do not cover every square cm of the viewgraph with something.

Text Foils (see also [Skeleton Foils](#))

There is a clear headline at the top

Rule of Thumb: 25 words or 7 lines per topic:

- This will keep it readable.
- It forces you to be concise.

Lower and upper case letters:

- Simple! UPPER CASE LETTERS ARE HARDER TO READ

Telegram style is what's needed:

- Keywords instead of sentences.
- Complete sentences will tempt you to read them out loud.
- *Never ever* read out loud what is written on your foil. Your audience will not include analphabets!

One thought per topic!

Structure and *emphasize* with colour.

- But don't get too *colourful*: Two to three colours are sufficient.
- Mark *essentials* with colour.

Have essentials, if possible, at the top or bottom of the foil.

- This goes against common feeling, but is a well known composition principle in art.

Tables

There is a clear headline at the top!

Orders of magnitude and units.

- Try to have units "understandable" to your audience. For physicists and material scientists, e.g., use **eV/atom**, for chemists **kJ/mol** for the same thing.
- Give no more than three digits if possible
- Units and multipliers (e.g. "**• 10⁶**") belong in the *heading* of rows and columns.

Structure of a table

- Vertical structures are easier to comprehend.
- Keep the decimal points aligned.
- Use the structures your audience knows and expects.

Emphasizing some points

- Mark *directly* with **bold letters** or in *colour* whatever you want to draw attention at upon presenting the table
- Underline or mark *during the presentation* when you want to make a point that is not directly obvious.

Diagrams

There is a clear headline at the top!

Quantity of information - some general rules

Of course, in scientific presentations you may have good reasons not to stick to these rules. But make sure, they are *really* good reasons. Not having enough time or energy to redraw an old diagram with too many graphs is not a good reason!

- 15 - 20 data points - no more!
- At most 4 graphs in one coordinate system
- No more than 3 columns in column diagrams
- At most 6 sectors in cake diagrams

Lines and areas

- Use strong primary colours for lines and pastels for areas.
- Make your graphs in strong lines, differentiate by strong colours.
- If colour is not available, differentiate by thick and thin lines, not by point-dash sequences.

[Previous page](#)

[Next page](#)

Viewgraphs - Some General Rules

Advanced

Format

- Mostly, the "*landscape*" format is preferable - it makes better use of available space. It also leaves more room for comparisons, which are much better in the left-right visualization than in the top-bottom format
- There are, however, many exceptions where the "*portrait*" format is better. Think about it before you start writing.

Make it lively!

- Work with your viewgraph! Add something; underline, circle,...
- But be careful with "striptease", i.e. covering parts which you uncover bit by bit. It is usually not a good idea - your audience feels patronized.

Use flip-overs

- It is often a very good idea to develop a point by putting a second viewgraph on top (a flip-over) while developing an idea. Use at most two flip-overs.
- Very important: Put them *firmly together* with scotch tape, so the flip-over will be exactly in place. Be aware of free standing projectors! Your flip-over foil will hang down and the whole viewgraph may slide off! Have a roll of scotch tape with you!

Take your viewgraphs *out of their jackets*! After all, you put a lot of work into them. Why obscuring the brilliance of your colors and so on by adding unnecessary absorption by a jacket?

9. Using Media

Presenting Viewgraphs for Illustrations - The 5 Step Procedure

1. Announce

- Get your audience in the mood by announcing the next step without preempting the information
- Example: "How would that look in the new system?"

2. Show

- Make a small pause while presenting the viewgraph
- This gives the audience time to look at the viewgraph and it gives you time to think about what you are going to say.

3. Explain

- Go through everything on the viewgraph in telegram style, and always, *always!!!* name the *axis'* and, in micrographs, the *scale*!
- But never, *never!!!* just read what is written on the viewgraph. This is the deepest insult you can hurl at your audience. They all can read it much faster themselves than you can spell it out! And if your viewgraph is clear, they will even understand it.
- If you believe you have to read it because the print is so small that the audience can not read it - don't worry, your audience after the second viewgraph of this kind, will neither read it nor listen to you, and your boss is going to fire you anyway. So the impression you left with that presentation doesn't matter anymore.

4. Meaning

- When the audience starts looking at you again, they are asking you a question: What does it mean?
- Answer that question! If this question does not come up, you must have presented your last and finishing viewgraph (for the whole presentation or for a main chapter), or you presented a meaningless viewgraph.

5. Resumée

- Give a short conclusion

⚡ You must at least allow **1 minute** per viewgraph! However, for complicated scientific stuff (formulas, several graphs, ...) **3 minutes** are more like it.

- There are exceptions, of course. You may insert a quicky just to illustrate **one** point ("The fundamental difference between the materials silicone and silicon, though unknown to the media professions, is **best illustrated** by the major application of silicone, which has nothing to do with the silicon used for **micro**electronics, but with rather **macro**scopic applications leading to soft products not easily confused with hard and brittle chips.").
In this case your illustration needs only **20** seconds to get the point across.

⚡ Keep the room as bright as possible! Don't turn off all lights - the ones next to the screen should be enough.. In a dark room, people get **very tired**!

How to Explain Visualized Topics

Wrong

- "I have prepared a foil for this..."
- "This is meant to explain how the space charge region.."
- "I hope you recognize that the symbol with the means..."
- On the **x**-axis I have inserted the times from **1968** to ..
- "As you can see immediately..."

Good

- "What is the situation now?" (present the foil)
- "The space charge region in this case..."
- (Pointing at the symbol) "This symbol means..."
- (Pointing at the axis) " the years from **1968**"
- You can see from this that..."

⚡ The differences seem to be rather small - because we are talking the emotional channel here! And in this part of the communication channel, the perception can be very different depending on how you explain your visualization. Don't insult your audience by pointing out the obvious or implicitly assuming that everybody already knows everything!

[Previous page](#)

[Next page](#)

10. The Talk

General Rules

Advanced

- ▶ Talk *loudly* and *clearly*.
 - Talking loud does not mean you have to scream.
 - Do not mumble. Rather than going off into silent obscurity if you are stuck, start the sentence again from the beginning.
 - Speak slowly. Not *e x t r e m e l y s l o w l y*, but a little bit under your normal rate.
- ▶ Make small *breaks* (5 seconds is already a long break - try it!)
 - The more important the statement, the longer the break after it!
 - Give the audience a little time to absorb your immortal words!
- ▶ Use *short sentences*
 - Rule of thumb: **15** words per sentence is about right
 - **Subordinate clauses** should be subordinate! Use at most one subordinate clause per **main clause**
 - Finish your sentence! Only the real masters of oral presentations can jump from one subject to the next with all kinds of circumlocutions in between and without ever finishing a sentence and, by doing this while in addition not sticking to all of the other subjects, which, by the way, are of course only for beginners anyway, who easily get confused which is not what we want to have here; but anyway, as I was going to say, your audience will be enraptured - or so we hope!
- ▶ *You* instead of *me*.
 - Address your audience, not yourself. Say "What *you* see here..." and not "*I* will show you..."
 - But say "*me*" whenever it is necessary. "*I* am of the opinion that..." instead of "One is lead to believe..."
- ▶ Use verbs instead of nouns. Compare:
 - The *preparation* of *topics* for the *purpose* of *presentation* requires from the *speaker* the *knowledge* of his *field of science* and in addition a *consideration* of the *findings* concerning *ways* of *information transfer*
 - If you want to make a good *presentation* you should know your *subject*. You should also present your *information* in such a *way* that your *audience* likes it.
 - We all know people who speak and write in nouns only. We do not like these people.
- ▶ Use the active and not the passive mode!
- ▶ Emphasize intentionally the main points. There are many tricks.
 - **Raise your voice** - or lower your voice - either way you get attention
 - Use breaks as mentioned before.
 - Repeat the sentence after a short break.
- ▶ Avoid meaningless noises to gain time!
 - The most common meaningless noise is the good old "hmm", or "hein", but there are many more
 - Fill words or clauses as, e.g. "in principle", "so to speak", "in other words", "as I have said before", ...

11. Remember

⚡ If you get perfect in all that was pointed out, nothing will stop you from becoming a politician or TV person.

⚡ However, if you want to become a good engineer or scientist, you must add one more thing:

● **You must know what you are talking about!**

[Previous page](#)

[Next page \(Checklist\)](#)

Check List for Oral Presentations

General Structure

- **Message**
Was the message clear? Did you clearly understand what the speakers wanted to get across?
- **Construction**
Was the construction in a logical order? Could you follow effortlessly from the introduction to the end?
- **Clarity**
Was the presentation graphical and demonstrative with good examples? Or very abstract?
- **Target Group Orientation**
Was the level below or above your level of knowledge? Did you learn something new or was it old stuff? Or did you not understand most of what was said because it was too sophisticated?

Content

- **General**
Was the content adequate? Too little or too much data, facts and theory?
- **Depth**
Were the speakers well prepared? Did they do their homework or just rehashed some article?
- **Knowledge**
Did the speakers understand their subject (i.e. they conveyed the impression that they know much more about it)? Or is what they said all they have?

Rhetoric

- **Dynamics and Modulation**
Was it lively? Or monotonous?
- **Speed and Pauses**

Speaking too slow or too fast? Short breaks or incessant talking?
- **Sentence Structure**
Whole sentences not too long? Or rambling along without clear structure?
- **Idiosyncrasies**
Hems and ohs, gesticulation, Annoying or cute and attention getting?
- **English**
Acceptable international science English? Vocabulary and pronunciation OK?

Body Language

- **General Posture**
Self-assured and with "presence"? Or looked like he/she would rather have been somewhere else?
- **Eye Contact**
With audience? Or with outer space?
- **Movements**
Gesticulation and body movements? Or stiff as a stick?

Visualization

- **Vuegraphs Technical**
Overloaded or trivial, color or black-white, pleasing arrangements or ugly?
- **Vuegraphs Content**
Helpful or confusing; trivial or too complicated
- **Others**
Experiments, specimen, videos, ..

Multiple Choice Test

1.3.1 Conductivity 1

Start Multiple Choice

Multiple Choice Test

1.3.3 Conductivity 2

Start Multiple Choice


Multiple Choice Test



1.3.4 Hall effect

Start Multiple Choice

Exercise 1.3-1

Derive and Discuss numbers for μ

 Calculate numerical values for the *mobility* μ of some typical metals.

-  Take [typical \(metal\) values](#) for specific conductivity σ and concentrations of electrons n and then *calculate* typical numbers for the mobility μ - do not take the values from the table! If you do not understand the German link, use [this one](#).
-  Consider typical field strengths for metals by picking suitable current densities, and then derive typical values for the *drift velocity* v_D .



Solution

Exercise 1.3-2

Derive numbers for v_0 , v_D , τ , and l

Show that the claims made in the backbone text are actually true (for room temperature = **300 K**). Use the following equations taken from the backbone

- For the average velocity v_0 of a particle zooming around in the crystal:

$$v_0 = \left(\frac{3kT}{m} \right)^{1/2}$$

- For the mean time τ between scattering:

$$\tau = \frac{\sigma \cdot m}{n \cdot e^2}$$

- For the drift velocity v_D

$$v_D = - \frac{E \cdot e \cdot \tau}{m}$$

- For the minimal mean free path length l_{\min} obtained for $v_D = 0$:

$$l_{\min} = 2 \cdot v_0 \cdot \tau$$

Of course, you need numbers for the concentration n of the free carriers and for the specific conductivity σ

- Since we are essentially considering metals, you assume for a start that you have **1** free electron per atom if you want to find a number for n . Here are a few data needed for the calculation:

Atom	Density [kg · m ⁻³]	Atomic weight × 1,66 · 10 ⁻²⁷ kg	<u>Conductivity</u> σ × 10 ⁵ [Ω ⁻¹ · m ⁻¹]	Concentration Atoms n [m ⁻³] ???
<u>Na</u>	970	23	2.4	
<u>Cu</u>	8.920	64	5.9	
<u>Au</u>	19.300	197	4.5	


You may run into some trouble with the dimensions. Just look at conversions from, e.g. [eV] to [J], from Ω to V and A, and at the relations between Volt, Ampere, Watts and Joule.





Solution

Exercise 1.3-3

What does it take to build a 4 GhZ Microprocessor?

 A typical **MOS** transistor of **200x** (**x = 0 5**) vintage has a "gate length" (= distance between source and drain) of about **0.5 μm** and is run at about **3 V**

-  **1.)** What is the mobility the material (= semiconductor) must have? Discuss the result for known mobility values and consider the following points
 - Transistor speed = device speed ??
 - Mobility range for a given material ??
 - Could we have powerful **PCs** without micro- or nanotechnology ??
-  **2.)** How could you increase the speed for a given material
 - In principle?
 - Considering that there are limits. e.g. to field strengths?



Solution

Solution to Exercise 1.3-1

Derive and Discuss numbers for μ

First Task: Derive numbers for the mobility μ .

- First we need typical conductivities and electron densities in *metals*, which we can take from the [table in the link](#).
- At the same time we expand the table a bit

Material	ρ [Ω cm]	σ [Ω^{-1} cm $^{-1}$]	Density $d \times 10^3$ [kg m $^{-3}$]	Atomic weight w [$\times 1u = 1,66 \cdot 10^{-27}$ kg]	$n = d/w$ [m $^{-3}$]
Silver Ag	$1,6 \cdot 10^{-6}$	$6,2 \cdot 10^5$	10,49	107,9	$5,85 \cdot 10^{28}$
Copper Cu	$1,7 \cdot 10^{-6}$	$5,9 \cdot 10^5$	8,92	63,5	$8,46 \cdot 10^{28}$
Lead Pb	$21 \cdot 10^{-6}$	$4,8 \cdot 10^4$	11,34	207,2	$3,3 \cdot 10^{28}$

For the mobility μ we have [the equation](#)

$$\mu = \frac{\sigma}{q \cdot n}$$

With q = elementary charge = $1,60 \cdot 10^{-19}$ C we obtain, for example for μ_{Ag}

$$\mu_{Ag} = \frac{6,2 \cdot 10^5}{1,6 \cdot 10^{-19} \cdot 5,85 \cdot 10^{28}} \frac{m^3}{C \cdot \Omega \cdot cm} = 66,2 \frac{cm^2}{C \cdot \Omega}$$

The unit is a bit strange, but remembering that $[C] = [A \cdot s]$ and $[\Omega] = [V/A]$, we obtain

$$\mu_{Ag} = 66,2 \frac{cm^2}{Vs}$$

$$\mu_{Cu} = 43,6 \frac{cm^2}{Vs}$$

$$\mu_{Pb} = 9,1 \frac{cm^2}{Vs}$$

Second Task: Derive numbers for the drift velocity v_D by considering a reasonable field strength.

- The mobility μ was defined as

$$\mu = \frac{v_D}{E}$$

or

$$v_D = \mu \cdot E$$

So what is a reasonable field strength in a metal?

- Easy. Consider a cube with side length $l = 1 \text{ cm}$. Its resistance R is given by

$$R = \frac{\rho \cdot l}{F} = \rho \Omega$$

- A **Cu** or **Ag** cube thus would have a resistance of about $1,5 \cdot 10^{-6} \Omega$. Applying a voltage of **1 V**, or equivalently a field strength of **1 V/cm** thus produces a current of $I = U/R \approx 650\,000 \text{ A}$ or a current density $j = 650\,000 \text{ A/cm}^2$
- That seems to be an awfully large current. Yes, but it is the kind of current *density* encountered in integrated circuits! Think about it!
- Nevertheless, the wires in your house carry at most about **30 A** (above that the fuse blows) with a cross section of about **1 mm²**; so a reasonable current density is **3000 A/cm²**, which we will get for about $U = 1,5 \cdot 10^{-6} \Omega \cdot 3000 \text{ A} = 4,5 \text{ mV}$.
- For a rough estimate we then take a field strength of **5 mV/cm** and a mobility of **50 cm²/Vs** and obtain

$$v_D = 50 \cdot 5 \frac{\text{mV} \cdot \text{cm}^2}{\text{cm} \cdot \text{V} \cdot \text{s}} = 0,25 \frac{\text{cm}}{\text{s}} = 2,5 \frac{\text{mm}}{\text{s}}$$

That should come as some surprise! The electrons only have to move *very slowly on average* in the current direction (or rather, due to sign conventions, against it).

- Is that true, or did we make a mistake?
- It *is* true! However, it does *not* mean, that electrons will not run around like crazy inside the crystal, at very high speeds. It only means that their *net* movement in current anti-direction is very slow.
- Think of an single fly in a fly swarm. Even better [read the module](#) that discusses this analogy in detail. The flies are flying around at high speed like crazy - but the fly swarm is not going anywhere as long as it stays in place. There is then no drift velocity and no net fly current!

Solution to Exercise 1.3-2

Derive numbers for v_0 , v_D , τ , and I

First Task: Derive a number for v_0 (at room temperature). We have

$$v_0 = \left(\frac{3kT}{m} \right)^{1/2} = \left(\frac{8,6 \cdot 10^{-5} \cdot 300}{9,1 \cdot 10^{-31}} \frac{\text{eV} \cdot \text{K}}{\text{K} \cdot \text{kg}} \right)^{1/2} = 1,68 \cdot 10^{14} \cdot \left(\frac{\text{eV}}{\text{kg}} \right)^{1/2}$$

The dimension "square root of eV/kg " does not look so good - for a velocity we would like to have m/s . In looking at the energies we equated kinetic energy with the classical dimension $[\text{kg} \cdot \text{m}^2/\text{s}^2] = [\text{J}]$ with thermal energy kT expressed in $[\text{eV}]$. So let's convert eV to J (use the [link](#)) and see if that solves the problem. We have $1 \text{ eV} = 1,6 \cdot 10^{-19} \text{ J} = 1,6 \cdot 10^{-19} \text{ kg} \cdot \text{m}^2 \cdot \text{s}^{-2}$ which gives us

$$v_0 = 1,68 \cdot 10^{14} \cdot \left(\frac{1,6 \cdot 10^{-19} \text{ kg} \cdot \text{m}^2}{\text{kg} \cdot \text{s}^2} \right)^{1/2} = 5,31 \cdot 10^4 \text{ m/s} = 1,91 \cdot 10^5 \text{ km/hr}$$

Possibly a bit surprising - those electrons are no sluggards but move around rather fast. Anyway, we have shown that a value of $\approx 10^4 \text{ m/s}$ [as postulated in the backbone](#) is really OK.

Of course, for $T \rightarrow 0$, we would have $v_0 \rightarrow 0$ - which should worry us a bit ????. If instead of room temperature ($T = 300 \text{ K}$) we would go to let's say 1200 K , we would just double the average speed of the electrons.

Second Task: Derive a number for τ . We have

$$\tau = \frac{\sigma \cdot m}{n \cdot e^2}$$

First we need some number for the concentration of free electrons per m^3 . For that we complete the [table given](#), noting that for the number of atoms per m^3 we have to divide the density by the atomic weight.

Atom	Density [$\text{kg} \cdot \text{m}^{-3}$]	Atomic weight $\times 1,66 \cdot 10^{-27} \text{ kg}$	Conductivity σ $\times 10^5 [\Omega^{-1} \cdot \text{m}^{-1}]$	No. Atoms [m^{-3}] $\times 10^{28}$
Na	970	23	2,4	2,54
Cu	8.920	64	5,9	8,40
Au	19.300	197	4,5	5,90

So let's take $5 \cdot 10^{28} \text{ m}^{-3}$ as a good order of magnitude guess for the number of atoms in a m^3 , and for a first estimate some average value $\sigma = 5 \cdot 10^5 [\Omega^{-1} \cdot \text{m}^{-1}]$. We obtain

$$\tau = \frac{5 \cdot 10^5 \cdot 9,1 \cdot 10^{-31}}{5 \cdot 10^{28} \cdot (1,6 \cdot 10^{-19})^2} \frac{\text{kg} \cdot \text{m}^3}{\Omega \cdot \text{m} \cdot \text{A}^2 \cdot \text{s}^2} = 3,55 \cdot 10^{-16} \frac{\text{kg} \cdot \text{m}^2}{\text{V} \cdot \text{A} \cdot \text{s}^2}$$

Well, somehow the whole thing would look much better with the unit $[\text{s}]$. So let's see if we can remedy the situation.

Easy: Volts times Amperes equals **Watts** which is power, e.g. energy per time, with the unit $[\text{J} \cdot \text{s}^{-1}] = \text{kg} \cdot \text{m}^2 \cdot \text{s}^{-3}$. Insertion yields

$$\tau = 1,42 \cdot 10^{-28} \frac{\text{kg} \cdot \text{m}^2 \cdot \text{s}^3}{\text{kg} \cdot \text{m}^2 \cdot \text{s}^2} = 3,55 \cdot 10^{-16} \text{ s} = 0,35 \text{ fs}$$

The backbone thus is right again. The scattering time is in the order of [femtosecond](#) which is a short time indeed. Since all variables enter the equation linearly, looking at somewhat other carrier densities (e.g. more than 1 electron per atom) or conductivities does not really change the general picture very much.

Third Task: Derive a number for v_D . We have (for a field strength $E = 100 \text{ V/m} = 1 \text{ V/cm}$)

$$\begin{aligned} |v_D| &= \frac{E \cdot e \cdot \tau}{m} = \frac{100 \cdot 1,6 \cdot 10^{-19} \cdot 3,55 \cdot 10^{-16}}{9,1 \cdot 10^{-31}} \frac{\text{V} \cdot \text{C} \cdot \text{s}}{\text{m} \cdot \text{kg}} = 6,24 \cdot 10^{-3} \frac{\text{V} \cdot \text{A} \cdot \text{s}^2}{\text{m} \cdot \text{kg}} \\ &= 6,24 \cdot 10^{-3} \frac{\text{kg} \cdot \text{m}^2 \cdot \text{s}^2}{\text{m} \cdot \text{kg} \cdot \text{s}^3} = 6,24 \cdot 10^{-3} \text{ m/s} = 6,24 \text{ mm/s} \end{aligned}$$

This is somewhat larger than the [value given in the backbone text](#).

- However - a field strength of **1 V/cm** applied to a **metal** is huge! Think about the current density j you would get if you apply **1 V** to a piece of metal **1 cm** thick.
- It is actually $j = \sigma \cdot E = 5 \cdot 10^7 [\Omega^{-1} \cdot \text{m}^{-1}] \cdot 100 \text{ V/m} = 5 \cdot 10^9 \text{ A/m}^2 = 5 \cdot 10^5 \text{ A/cm}^2$!
- For a more "reasonable" current density of **10^3 A/cm^2** we have to reduce E hundredfold and then end up with $|v_D| = \mathbf{0,0624 \text{ mm/s}}$ - and that is slow indeed!

Fourth Task: Derive a number for I . We have

$$I_{\min} = 2 \cdot v_0 \cdot \tau = 2 \cdot 5,31 \cdot 10^4 \cdot 3,55 \cdot 10^{-16} \text{ m} = 3,77 \cdot 10^7 \text{ m} = 0,0377 \text{ nm}$$

- Right again! If we add the comparatively miniscule v_D , nothing would change. Decreasing the temperature would lower I to eventually zero, or more precisely, to $2 \cdot v_D \cdot \tau$ and thus to a value far smaller than an atom..

Solution to Exercise 1.3-3

What does it take to build a 4 GhZ Microprocessor?

First Task: What is the mobility the material (= semiconductor) must have? Discuss the result in considering the following points

- Transistor speed = device speed ???
- Mobility range for a given material ??
- Could we have powerful **PCs** without micro- or nanotechnology ??

The [essential equation](#) is

$$t_{SD} = \frac{I_{SD}^2}{\mu \cdot U_{SD}} \approx \frac{1}{f_{max}}$$

The necessary mobility thus is given by

$$\mu = \frac{I_{SD}^2}{t_{SD} \cdot U_{SD}} = \frac{f_{max} \cdot I_{SD}^2}{U_{SD}} = \frac{4 \cdot 10^9 \cdot 2.5 \cdot 10^{-13}}{3} \cdot \frac{m^2}{s \cdot V} = 0.33 \cdot 10^{-3} \frac{m^2}{s \cdot V} = 3.3 \frac{cm^2}{s \cdot V}$$

What is the mobility of typical semiconductors? Finding values in the Net is not too difficult; if you just turn to the Hyperscript "[Semiconductors](#)" you should find [this link](#)

- Well, all "useful" semiconductors seem to be OK, their mobilities are much larger than what we need. But perhaps we are a little naïve?
- Yes, we are! If a device combining some **10.000.000** transistors is to have a limit frequency of **4 Ghz**, an individual transistor "obviously" must be much faster. If you don't see the obvious, think about the routing of many letters by the mail through a few million post offices (with different routes for every letter) and compare the individual and (average) total processing times.
- Bearing this in mind, mobilities of about a factor of **100** larger than the one we calculated do not look all that good anymore!

The mobility table in the link shows large variations in mobility for a given material - obviously μ is not really a material constant but somehow depends on the detailed structure.

- We do not need to understand the intricacies of that table - [we already know](#) that μ is directly proportional to the mean free path length l and thus somehow inversely proportional to defect densities.
- It is very clear, then, that for high-speed devices we need rather perfect crystals! So let's try to have single crystals, with no dislocations (or at least only small densities, meaning that the crystal must *never* plastically), and the minimum number of extrinsic and intrinsic point defects.
- Quite clear - but do you see the *intrinsic* problem? A more or less perfect crystal is *not* a device! To make a device from a crystal, we must do something to the crystal. And whatever you do to a *perfect* crystal - the result can only be a less perfect crystal!
- In other words: Making a device means to start with very good crystals and only induce the minimum of defects that is absolutely necessary.

Could we have **4 GHz** without microelectronics?

- Well, take for I_{SD} a value **100** times larger, and your highest frequency will be **10.000** times smaller - **400 kHz** in the example. Of course, the **4 GHz** of modern processors is not only determined by mobility values of the materials used, but the argument is nevertheless valid.
- So, without microelectronics (or by now nanoelectronics) life would be much different, because you can just about forget everything you do as a direct (and indirect!) present-day "user" of electronics. But would it be worse? The answer is a definite: Yes - it would be worse! Trust me - I have been there! It's not that long ago that **400 kHz** was considered a pretty high frequency.

Second Task: How could you increase the speed for a given material

- In principal
- Considering that there limits. e.g. to field strength

In principal it is simple: Make I_{SD} smaller and / or U_{SD} larger.

It is so simple, that you now should wonder, why it's not done immediately? Why not make a **40 GHz** or **400 GHz** microprocessor now - always, of course, only as far as it concerns the mobility?

Well, there are limits that are not so easily overcome. To name just two:

Things are structured by "painting" with light. And just as much as you can't make a line thinner than the size of your brush or pencil, you can't make structures smaller than the wavelength of the light you use, which is in the **0.5 μm** range.

Funny coincidence to the I_{SD} we used, don't you think so?

OK, so we increase the voltage; let's say from **3 V** to **300 V**.

This increases the field strength from $3/5 \cdot 10^5 \text{ V/cm}$ to $3/5 \cdot 10^7 \text{ V/cm}$ or **600.000 V/mm**.

In other words: A **1 mm** thick layer of your material should be able to isolate a high-voltage cable carrying **600.000 V**. Seems a bit strange, given the fact that they still hang lousy **300.000 V** cables high up on poles to have many meters of air (a very good insulator) because otherwise you would have to use many **cm** of some really good insulating solid.

To put it simple: no material withstands field strength of more than **10 MV/cm** (give or take a few **MV**). If you try to exceed that value, you will get interesting and very loud fire works. Whenever mother nature tries it, we call it a thunderstorm.

And only a few very good *insulators* will even come close to that number. Semiconductors, not being insulators, by necessity, can take far less. Our **60.000 V/cm** are pretty much the limit. So forget about higher voltages, too.

Does this mean **4 GHz** is the end of the line?

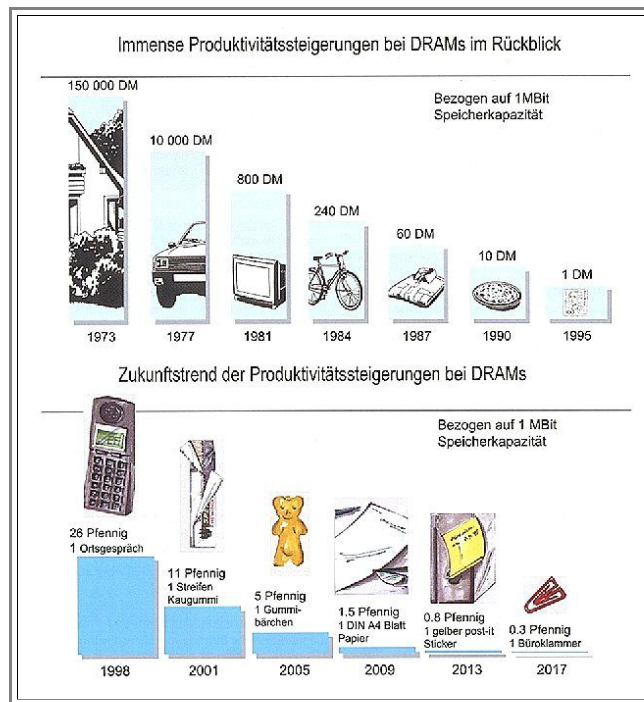
No it's not. It just means it is not easy to go beyond. It takes a lot of knowledge, understanding, and skills to make existing devices "better". It takes highly qualified engineers and scientists to do the job. It takes what you will be in a few more years if you keep to it!

Example for Illustrating Tables

Here is the boring table

Year	1973	77	81	84	87	90	95	98	01	05	09	13	17
Price for 1 Mbit of DRAM memory (DM)	150.000	10.000	800	240	60	10	1	0,26	0,11	0,05	0,014	0,008	0,003

And here is an interesting way for illustrating the same data



Example for Skeleton Foil

Sorry, it's in German; but the idea should be clear

Illustration

The original had a little hand colouring; the "markers" (the circles, stars, cartouches) had a specific colour and a specific meaning

- Circles denote headlines or are main topics
- Stars are major points and facts
- Cartouches indicate actions:
 - "Folie 1" means that here an illustrative viewgraph will be shown; after that this foil comes back
 - "Muster" means that a specimen is passed around or hold up to see.

Otherwise the speaker talks freely to the topics on the foil, following his own lead.

SIEMENS
4 M - DRAM

● Das MEGA - Projekt

★ Aufholen bei 1 M - Technologie } gemeinsam
 Gleichziehen bei 4 M - Technologie } mit Philips

★ ca. 1200 Mannjahre
 ca. 1,5 GDM Kosten
 ca. 250 Neueinstellungen in München P
 ca. 350 neue Arbeitsplätze in Regensburg

★ zwei 6" - Linien und Labor / Bürobauten in Mch P

Folie 1 Mikroelektronik - Zentrum in Mch P

★ Zeitplan 4 M - DRAM

"Erstes Silizium"	----->	3/87	}	Ziele
Kundenmuster	----->	1988	}	erreicht oder erreichbar
Fertigungsanlauf	----->	1989		

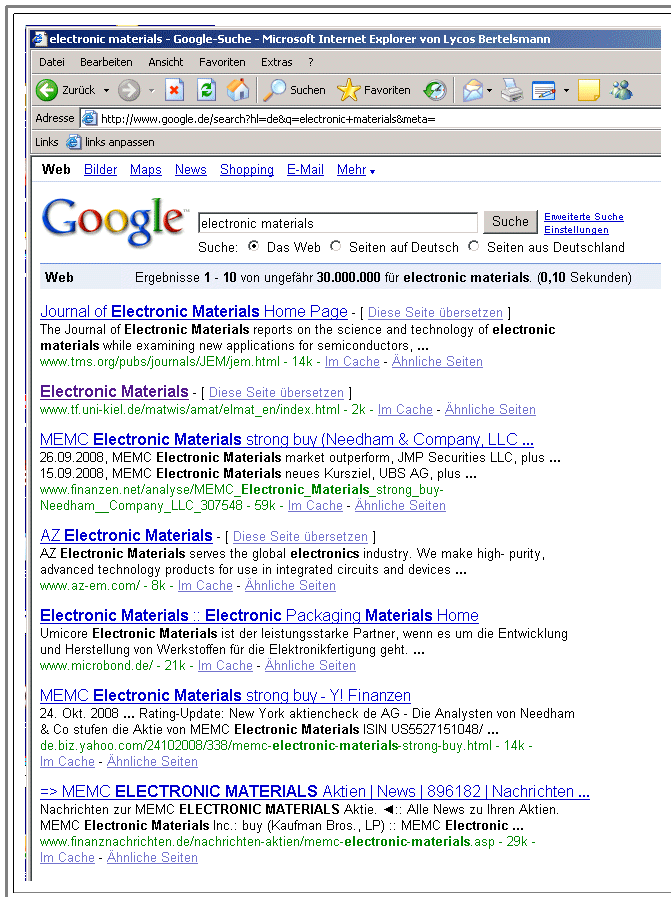
Muster Erste funktionsfähige 4 M - DRAM's

1. Einleitung
HLT 3 / Fö

What are "Electronic Materials"

Ask Google, and you get this answer:

Illustration



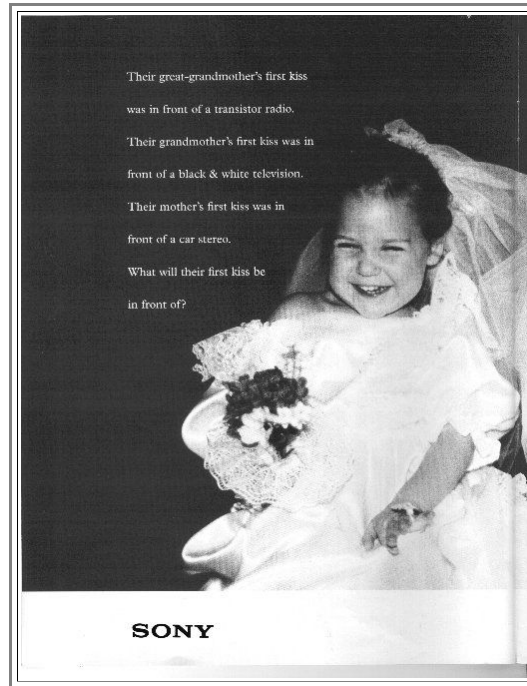
People forget

Illustration

People do forget things. But often it is just an excuse for major blunders, used if you simply didn't *know* the issue, didn't feel like *looking it up*, or just plain *guessed*. If you forgot some of the topics needed for the course:

● In engineering, guessing is generally a bad idea. In case of doubt: *Look it up!*

● Your mistakes are potentially dangerous. You will not just make a fool out of yourself like the "creative" genius (In Germany all advertising people call themselves "Kreative") who did the add shown below. I wonder if (s)he ever noticed.



● We do know, of course, that people behind advertising quite generally have a somewhat disturbed relation to the truth. But here we see that they are also very generous when it comes to recent technical history.

● Above, the left part of a two-page add is shown that appeared for quite some time in all major magazine in the **US** (around **1998**). If one assumes that the **SONY** people look at their adds before they are printed, this beauty demonstrates very nicely that the really fundamental human achievements in this century did not make a deep impression on our "creative thinkers".

● The great-grandmother alluded to above lived about **80** years ago, i.e. around **1910 - 1920**. That the quantum theory, the base of transistors was not yet quite invented, must be seen as an irrelevant detail. Unfortunately, however, the transistor itself was not invented before **1947** (by John **Bardeen**, Walter **Brattain** and William **Shockley**) - tough luck! "Transistors", meaning battery powered mobile radios, only hit the market in the sixties. Great-grandma, by the way, did not listen to the radio at all - it didn't exist then.

● Grandmother too, hardly enjoyed her first kiss in front of the black-and-white **TV**. It just barely existed, but not many people could afford it or lived in areas where there was actual broadcasting.

● Mother, that much could be true, may have listened to the stereo in the car while making-out; car radios and stereos did exist about **20 - 30** years back.

● Colour **TV** it appears, was not so impressive as to be mentioned. And vacuum tubes obviously never existed at all in the world of our "creative" geniuses.

Silicone vs. Silicon

The fundamental difference between the materials **silicone** and *silicon*, though unknown to the media professionals, is best illustrated by the major application of *silicone*, which has nothing to do with the *silicon* used for *micro*electronics, but with rather *macro*scopic applications leading to soft products not easily confused with hard and brittle chips.

Illustration



Requirements for Chip Metallization

The metal lines connecting transistors or other components on a **Si** chip must meet many, partially conflicting, requirements. Below is a list, including some materials that do *not* meet the particular requirement very well.

Can you guess the winner?

Advanced

Desired Property	Materials not meeting requirement
Very good conductivity	All but Ag, Cu
High eutectic temperature with Si (> 800 °C would be good)	Au, Pd, Al, Mg
Low diffusivity in Si	Cu, Ni, Li
Low oxidation rate; stable oxide	Refr. Metals, Mg, Fe, Cu, Ag
High melting point	Al, Mg, Cu
Minimal interaction with Si substrate	Pt, Pd, Rh, V, Ni, Mo, Cr (form silicides easily)
Minimal interaction with poly Si	Same as above
No interaction with SiO₂	Hf, Zr, Ti, Ta, Nb, V, Mg, Al
But must stick well to SiO₂	?
Must also comply with other substrates, e.g. TiN	? (see example for Al)
Chemical stability, especially in HF environments	Fe, Co, Ni, Cu, Mg, Al
Easy structuring	Pt, Pd, Ni, Co, Au
Electromigration resistant	Al, Cu
.... and many more,...	

The winner is: **Aluminum** (with <1% of **Si** and **Cu** added).

Al, in fact, is pretty bad - but all others are worse!

Presently (**2001**) a switch to **Cu** takes place (the better conductivity is definitely needed). The industry will pay several **10⁹** Dollars to develop the new material technology and change the production facilities.

Do Not Forget the Temperature Dependence of the Specific Resistivity!

Advanced

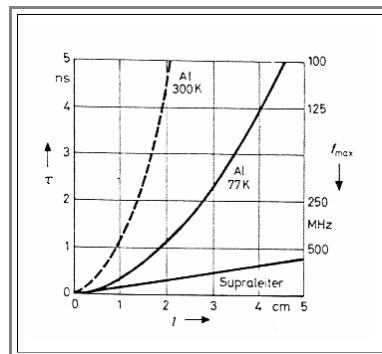
- The discovery of high temperature superconductors in **1986** immediately lead to proposals to use these materials for interconnects on chips instead of the **Al** that was common then (and for about **15** more years).
- The reason was that the finite resistivity of **Al** together with parasitic capacitances (e.g. between two conducting lines on a chip) limits the maximum frequency to

$$f_{\max} = \frac{1}{R \cdot C}$$

- With **R** = resistance of the longest connection line on the chip and **C** = parasitic capacitance "seen" by this line.
- For **R = 0 Ω** as we have it for a superconductor, the maximum frequency is no longer limited by **R · C**, no matter how large the parasitic capacitances are. Instead, the limit comes from $f_{\max} = (L \cdot C)^{-1/2}$ with **L** = inductance of the line, and this is just another way of saying that the signal propagation is limited by the speed of light.

$$f_{\max} = (L \cdot C)^{-1/2}$$

- Given the resistivity of **Al** (at room temperature!), a sizeable advantage was seen for the integrated circuits then envisioned.
- However, comparing the performance of a chip run with **Al** at room temperature to a chip run at liquid **N₂** temperature (**77 K**), is not the right comparison. After all, you can cool down the conventional chip, too - and that will decrease **R_{Al}** by a factor of **6 - 8**.
- The comparison then is quite different. The graph shows the minimum switching time $\tau = 1/f_{\max}$ as a function of the length of a standard interconnect line about **1 μm²** cross section.
- Whereas superconductors would already make an interesting difference for lengths of a few **mm** (typical line length) in the **wrong** comparison, the correct comparison only shows an advantage for about **1 cm** and larger - line lengths easily avoided by clever design.



The Li Ion Battery

The Issue in Short

Advanced

- Presently (Oct. **2009**) the world at large is in the throes of the financial crisis; before and after it was (and will be) shaken by the **energy crisis**. The articles "[Terawatt challenge](#)" and "[Powering the planet](#)" provide some background of particular interest to Material Science and Engineering students.
- Every budding engineer knows, of course, that automobiles in the near future will run **electrically**, powered by **batteries** - at least this is the unavoidable conclusion if one just reads a newspaper occasionally or listens to news.
 - Is that true? Are we all going to drive around in electrical cars (or at least serious hybrids) in a few years that have about the same cost / performance relationship than present gasoline powered cars? This module will provide a few essentials for pondering this question.
- First, we realize that the energy source in a future car will be a **Li ion battery**? Why is that? We will find out by recalling some basic facts about cars and doing a little quiz.
- Before you read on, you should try to figure out the answers yourself by doing the extremely simple but highly illuminating quizzes (far too simple to be called "exercise") provided in the links.

Quiz 1

How much energy is contained in **1 liter** of gasoline?
Find out by using data you know about your car.

- What you could get is shown in the following table.

Known Property	SI units	Stupid units
Mileage	10 km/l	23.5 miles/gallon
Range with full tank	500 km	311 miles
Engine Power P	100 kW	135 PS or horse powers
Average speed in going 500 km	120 km/h	74.6 miles/h
Average capacity of engine used	50 %	
Average efficiency η of engine used	0,3 or 33%	
Time t to cover 500 km	4.17 h	
Energy E consumed = $0,5Pt$	208 kWh	
Energy E_G contained in 1 l of gasoline = $E/50 \text{ l} \cdot \eta$	12.6 kWh/l	

- Not bad! The "official" number is $E_G \approx 10 \text{ kWh/l}$.
- The " \approx " sign alludes to the (small) differences between premium, diesel, and so on. Who cares. We are only concerned about orders of magnitude here.
- Is a specific energy of around **10 kWh /l** a *large* or *small* number? Find out yourself in the next quiz:

Quiz 2

How large or small are **10 kWh /l** - in comparison to known energy (densities).

- Here are some answers: **1 kWh** energy is stored or used up in

- 0,1 l gasoline
- Large (**85 Ah**) truck battery.
- 0,25 kg dry wood
- 7,3 t H_2O in a reservoir with **50 m** height difference.
- A weight of **367 t** lifted to **1 m**.
- 9,5 l brought to a boil (**100 °C**) from **10°C**.

- **100 h** of bicycling at **100 W** uses up the **10 kWh**, moves you about **2.000 km**, and works off about **17 Big Mac's**.

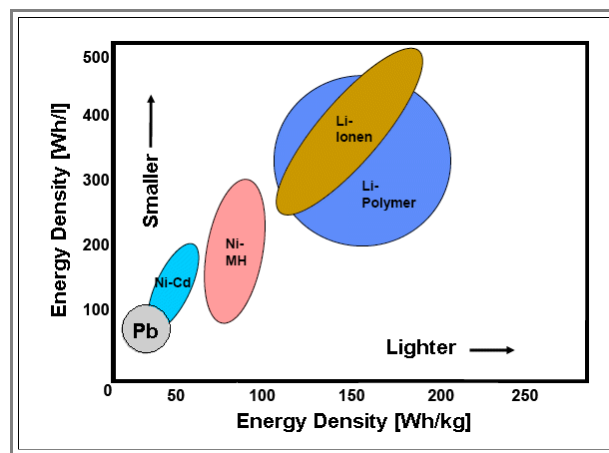
In other words: The chemical energy contained in gasoline is *absolutely huge* compared to typical mechanical energies we can directly relate to! It is, by the way, larger than the energy stored in **1 kg** of a solid high explosive.

So what is the energy density of a decent battery? Well - it depends. Do you mean energy *per liter* or *per kilogram*? There is a big difference from the view point of the user.

Quiz 3

This and That

- If you think about electric cars, the **weight** of the battery is what you are concerned with. The volume is not all that important. You have plenty of space because you do no longer need a bulky engine, transmission, drive train and so on. Electric motors are much smaller than comparable gasoline counterparts.
 - If you want to power your submarine, it's the the other way around. If the batteries are too light weight, you have trouble to "sink" it. If you want to power your cell phone, it's volume you are concerned about. If you want to store huge amounts of energy (e.g. to get through the night if your major energy source are solar cells), weight and volume are not confining - the price is!
- Let's look at both specific energy densities in one diagram



- It looks like the best we can come up with is at present (Oct. 2009) the **Li ion** battery with a specific capacity around **0.5 kWh per liter** or **0,15 kWh per kg**. In other words: **1 l** of gasoline beats the best battery we have by a factor of **20** with respect to the specific volume and by more than a factor of **100** with regard to the specific weight.
 - OK - that simply means we have to find room for a **Li ion** battery with a volume of **1000 l** or **1m x 1m x 10 cm** to run a car. That should be possible in a decent sized car.
 - Unfortunately, if we look at the figure, the weight of such a battery would be around **3 000 kg** or so (and the price would be totally prohibitive). Twice the weight of your old-fashioned standard car!
- The conclusion is easy:

If cars are to be powered by batteries in the near future, we need to improve batteries at at least by a factor of 5 or so with respect to the energy density.

As a material scientist you now wonder what exactly determines the energy density of a battery. For finding out we need to look at the basic working principle of a battery.

Working Principle of a Battery

The energy stored in a battery is chemical energy, i.e. the energy released by a chemical reaction. We will only consider **Li** based batteries here, so we have a reaction [of the type](#):



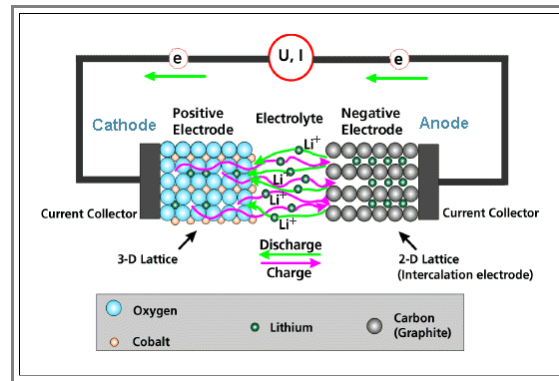
So take a piece of **Li** and a "piece" of **F**, separate it by an electrolyte that cannot pass electrons but only ions, put it into a box with contacts to the **Li** and the **F** - you have a battery.

It's not so easy, of course. The general principle of forcing the electrons to go through an outer circuit from one reaction partner to the other one, while the ions move through the electrolyte is certainly correct. But:

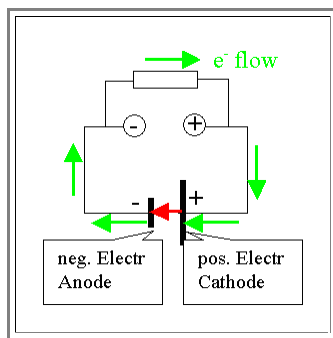
1. How do you keep and contact an (extremely corrosive) gas like F_2 ?
2. How do you keep your **Li** from reacting with the oxygen in the air?
3. Will the reaction actually take place? Will a piece of metallic **Li** in a F_2 atmosphere start to react without being "triggered" somehow (like $\text{H}_2 + \text{O}_2$)?
4. Can you reverse the reaction by running a current through your battery, i.e. charge it again?
5. and so on. and so forth.

Making a real battery with "**Li ions**" is not all that easy - Point 2 will always be a problem, for example, demanding airtight sealing and a lot of security features.

So let's look at the working principle of a **Li ion battery** in a very general way.



We have a **negative electrode**, also called **anode**, where **Li** is present as a "metal".



- The term **negative electrode** is clear. The **Li** atoms incorporated or **intercalated** in the electrode material must leave an electron back so that they can move as Li^+ ion through the electrolyte if the battery is discharged and thus supplying energy to the outside world.
- But why is it also called **anode**? Isn't an anode always the positive pole towards which the electrons flow?
- Yeah - but only **outside** the battery as shown on the diagram. **Inside** the battery the electrons (or negative charges) must flow from **plus to minus**. The battery, after all, is an electron pump that moves the electrons "up" to a high energy level from which they can "flow down" towards the positive pole.
- Of course, it's not electrons flowing from plus to minus inside the battery but positively charged Li^+ ions in our case. But if you replace the battery by a generator, your electrons flow indeed in the "wrong" direction inside the generator.
- The general definition of an anode is therefore: An anode accepts electrons; it provides for a general oxidation reaction. The polarity is positive seen from an external user and negative as seen from inside a generator. In the picture you see that the internal electron (or negative charge) current does indeed flow into the anode (red arrows).

We have a **positive electrode** also called **cathode** (now we know why) where **Li** reacts to whatever and is incorporated as Li^+ .

The chemical reaction to "whatever" - e.g. LiCoPO_4 - provides the energy that drives the whole process.

Between the electrodes is an **electrolyte** that allows Li^+ ions to pass but does not conduct electrons. That's why we have a battery. The electrons can only get from the anode to the cathode by flowing through electron conductors in the external circuit.

The two electrodes must be intimately connected to some **current collector** (a metal like **Cu**) that conducts the electrical current out of the battery to the terminals.

This is not a simple issue! If you want to run a **100 kW** motor with a **3 V** battery, you are going to draw **100 000/3 A = 33 333 A** which is a lot of current. Even if you switch battery cells in series to achieve **300 V**, you still run **333 A** through the current collectors and wires.

What kind of properties are we asking for concerning the electrode materials? It's a long list; most important for us at this point are

General Properties:

- Specific capacity or how much **Li** (in **kg**) can you incorporate in **1 kg** or **1 l** of the the electrode material.
- Electrochemical potential.
- Conductivity; connection to current collector.
- Long term stability; survives how many charge / discharge cycles?
- Hazard potential (does it explode / burn if exposed to air?).
- Price.
- Ecological topics.

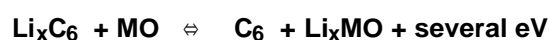
Negative Electrode/Anode

- In principle one could use **Li** metal. In practice, however, one cannot, for many reasons. Consider just recharging after all your **Li** has been used up and is now incorporated in the cathode. You now have a "hole" on your anode side. How do you get your **Li** back?
- We therefore use an anode material that ideally can incorporate a lot of **Li** easily, **and** can release most of that incorporated **Li** easily. In other words, it shouldn't take much energy to get it out again. That implies that no strong bonds between the anode material and the **Li** must develop.
- The standard material at present is **graphite**. **Li** atoms are simply "intercalated" between the hexagonal **C**-layers. Another, very interesting material for anodes is **Si**. We will come back to that.

Positive Electrode/Cathode Properties

- In principle same thing as above, except that we must produce a lot of energy whenever **Li⁺** is incorporated.
- What we use are metal oxides (**MO**) like **LiCoO₂**, **LiNiO₂**, **LiMn₂O₄**, **LiFePO₄**, **LiNi_xCo_yMn₂O₂**.

The over-all reaction then is like this:



All we do, in simple terms, is to "shuttle" **Li** back and forth between the two electrodes, gaining energy in one direction (discharge) and using energy (charge) in the other.

Capacity of a Battery

What determines the capacity of a given **Li** ion battery? Easy. We first do a quiz to get some ideas.

Quiz 4

Battery capacity

What we see is that about **90 g** of **Li** would be enough for a capacity of **1 kWh**. But the [diagram above](#) shows that presently we need at least **5.000 g** for that. How can that be?

- Well - besides the **Li**, you need the two electrodes that contain it, an electrolyte, current collectors, and a housing; not to mention security features.
- If we want to improve the gravimetric energy density by at least a factor of **5** [as we must](#), we realize that there is a lot of work out there for enterprising Material Scientists. What we have to work at are the electrodes (and everything else). The crucial question is: What is the specific capacity (in **mAh/g**) of an electrode. In other words: How much **Li** (in **g**) can I incorporate into **1 g** of the electrode material.
- The reference would be metallic **Li**, where one obviously can have "**1 g Li** in **1 g Li**" with a specific capacity of **3800 mAh/g**. It also takes a voltage of **0 V** to get **Li** "out", i.e. to drive the reaction **Li/Li⁺** because we use that as the zero point of the potential scale.
- The relevant voltages are shown for various reactions in the figure. The difference between the high end and the low end is what you get as the battery voltage. With a **Si/Li** anode and a **Li_xCoO₂** standard cathode, you should obtain around **3.5 V**.

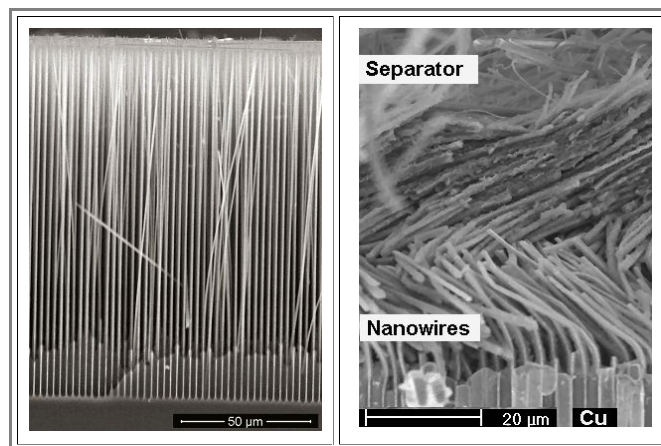
Let's look what else we have

Material	Specific capacity [mAh/g]	U vs. Li/Li ⁺ [V]	Comments
Graphite	330 - 370	0,1 - 0,6	Present standard anode
Li ₄ Ti ₅ O ₁₂	155	1,6	
Si	> 4 000	0,1 - 0,5	Volume change factor 4!!!
Li	3 800	0	Not practical

- Now you should be a bit surprised. How can you have more **Li** in **Si** than in **Li** itself?
- Yes - you can. By forming alloys like **LiSi**, **Li₁₂Si₇**, **Li₇Si₃**, **Li₁₃Si₄** and **Li₂₂Si₅**.
- So why don't we have **Si** anodes as a matter of course in our **Li** ion batteries?
- Because the volume of the **Si** expands by a factor of **4** when **Li** in high concentrations is incorporated. What that means is that your piece of **Si** anode will have fractured to fine dust if you load it just once with **Li**. This is obviously not practical.
 - Nevertheless it can be done. The keyword are "**nanowires**". Overcoming that problem by using **Si** nanowires instead of bulk **Si** is not only a story in itself but brings you to the heart of "hot" research in materials science and engineering as it is going on right now (Oct. **2009**).


The Si Nanowire Anode

- The pictures below show **Si** nanowire anodes. On the left we have a **Si** nanowire array as made by anodic pore etching and chemical over-etching: On the right we see a **Si** nanowire anode with a **Cu** backing after it has been used in a few charge and discharge cycles in a battery.



- It works! No destruction of the **Si** nanowires is visible.
- If you want to know more about this, refer to these two short papers:
 - The [Stanford paper](#) that started it all in **Jan 2008**.
 - The [Kiel paper](#) that makes it practical (and was co-authored by a student presently (fall term **2009/10**) taking this course).

Tunnelling effect

 This was actually a exercise in "Matwiss I". The [Link](#) takes you there.

Advanced

Thermoelectric effects

Seebeck Effect

Advanced

First of all, in considering thermoelectric effects, we have to realize that we are dealing with a *non*-equilibrium situation.

A general theory of non-equilibrium is beyond our means, suffice it to say that Lars **Onsager**, with a paper entitled "*Reciprocal relations in irreversible processes*" induced some fundamental insights as late as **1930**; he received the Nobel price for his contribution to non-equilibrium thermodynamics in **1968** - for chemistry, of all things.

However, what we should be aware of, is the essential statement of non-equilibrium theory:

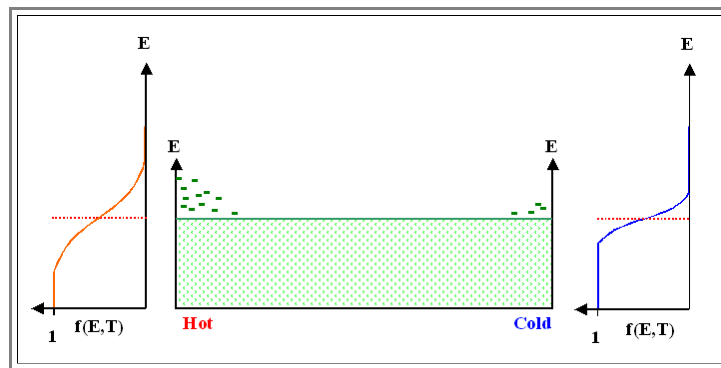
As long as there is no equilibrium, we always have *currents* of something trying to establish equilibrium by reducing a gradient in something else that is the actual cause of the non-equilibrium. A gradient in the electrical potential, e.g., causes our well-known electrical currents, and a gradient in a concentration causes diffusion currents.

But we must abstract even more, and consider things like entropy currents as well as all kinds of combinations of gradients and currents.

While Onsager discovered some quite general relations between gradients and currents, we will not delve into details here, but only look a bit more closely at what causes the Seebeck effect.

For that, we still treat the thermoelectric effect with equilibrium thermodynamics, simply assuming that *locally* we are not very far from equilibrium and thus can still use band structure models with a Fermi energy (which is only a well defined quantity for equilibrium) and resulting carrier distributions.

In the simplest possible case, what we will get for a long bar of metal, hot at one end and cold at the other, is something like this:



At the hot end, the Fermi distribution is "soft", and we have a noticeable concentration of electrons well above the Fermi energy. At the cold end, the Fermi distribution is sharp, and we have fewer electrons above the Fermi energy.

The drawing, of course, grossly exaggerates the real situation. Note also that the total concentration of electrons at both ends is the same - even so the drawing does not show this because the holes below the Fermi energy are not included.

Note too, that the Fermi energy is constant throughout the material (we neglect any possible effects of the temperature on the Fermi energy, as we have it, for example, in [doped semiconductors](#)).

As always, electrons go to where the energy is lower; the electrons would tend to move from the hot end to the cold end, thereby transporting energy and thus equilibrating the temperature eventually. Equilibrium, with a constant temperature everywhere will be achieved.

An equally valid alternative interpretation just looks at the concentration gradient of the electrons in energy space, which would automatically drive a kind diffusion current until the concentration (and thus the temperature) is equalized.

Yet another way of looking at it is to consider that the average momentum of the electrons at the hot end is larger than that of the electrons at the cold end. They would therefore "run away" faster (taking energy with them) than the electrons from the other end would "run in".

However, since we keep the *temperature difference* constant, all this *cannot happen*. We will have to maintain constant but different temperatures and therefore different energy distributions at both ends of the metal bar.

If nothing happens, we will lose the electrons with large momentum faster than we gain electrons with smaller momentum; and a temperature gradient cannot be maintained. The only way to change that, is to lower the potential at the hot end somewhat, i.e. make the ends positively charged, and to raise it at the cold end.

The potential difference must build up until it is large enough to exactly counteract the net loss of "hot" electrons due to momentum imbalance.

This is essentially the reason why we find a **thermoelectric voltage**.

- Note that the junction is not directly essential. However, if you just plug a wire from one material into your Voltmeter and heat up the middle part, leaving the two ends cold (and at the same temperature), your potential along the wire may change, but at the two ends you have the same potential, and it is the potential difference between the two ends you measure
- Plugging the hot end into your Voltmeter is a bit impractical, so you necessarily end up with a junction to some other material. The other material now will also have a hot end and a cold end, and thus develop a potential difference.
- Since the potential at the other end can only have **one** value, you will now get a potential difference between the **two** cold ends which depends, of course, somehow on the choice of materials.

Still, there is a potential difference between the hot and cold end of one piece of material, and even so it cannot be measured directly, we can measure it indirectly somehow and tabulate the values.

- We can do this, with somewhat more involved but similar reasoning not only for metals, but also for semiconductors. The table below gives some absolute values and shows that semiconductors are good candidates for actual thermocouples, because their **Seebeck voltage** is fairly large. The values are for about room temperature, or about **700 °C** for the last three materials

Material	Al	Cu	Ag	W	(Bi,Sb) ₂ Te ₃	Bi ₂ (Te,Se) ₃	ZnSb	InSb	Ge	TiO ₂
Seebeck voltage [μV/K] ($V_{\text{hot}} - V_{\text{cold}}$)	-0,20	+3,98	+3,68	+5,0	+195	-210	+220	-130	-210	-200

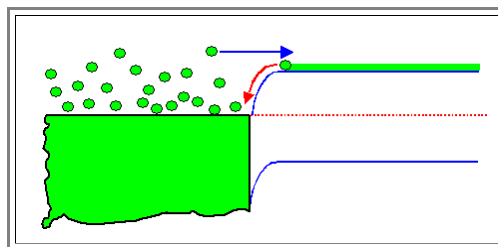
Surprise! There are positive (as expected) **and** negative values of the voltage. What does it mean?

- Simply that you are looking at **positively charged** carriers being responsible for the Seebeck effect - **holes**, in other words.
- Not so surprising for semiconductors, perhaps, but somewhat unexpected for **Al**. But, as we should know, conduction in **Al** relies heavily on holes, as evidenced, e.g., in its **positive Hall coefficient** while most other metals have a negative one

Peltier Effect

Understanding qualitatively the Seebeck effect does not help much to understand the Peltier effect.

- Again, let's look at some simple junction, this time an ohmic contact of a metal to a semiconductor.



- Shown is an equilibrium situation, where the Fermi energy is constant throughout, and the flow of electrons across the junction must be equal in both directions. Note that only the high-energy end electrons of the Fermi distribution in the metal make it across the junction, whereas all electrons of the semiconductor can flow into the metal.
- The electrons of the metal thus also transport some thermal energy out of the metal, but in equilibrium exactly the same amount is gained by the semiconductor electrons, which are high-energy electrons in the metal.

Now consider some external voltage driving some net current through the junction in either direction.

- If this current is an electron current flowing from the metal into the semiconductor, it **still** transports some thermal energy out of the metal, but since it is now much larger than the electron current flowing back, we have a **net transport of thermal energy** out of the metal, which therefore **must cool down** – and that's it; nothing happens to the semiconductor part due to these.
- If the current is reversed, the flow of thermal energy reverses, too, and now the metal at the contact heats up.
- It is conceivable then (also far from clear) that the total effect in terms of temperature change in the metal is proportional to the current I flowing.

Note, however, that as a completely independent process, you always have ohmic heating (or Joule heating) which is simply given by the total power P dumped into the system via

$$P = U \cdot I = R \cdot I^2$$

- with U = voltage applied, R = total series resistance of the system.
- Since this general heating of the whole device is proportional to I^2 , it can easily overwhelm any cooling effect that might be there.
- If you want to use the Peltier effect as an elegant way of cooling something, you must not only choose your materials very carefully, but also optimize your system design and working points.
- That this is possible is evidenced by the successful marketing of **Peltier cooling elements**, mostly for scientific applications. Here is a table with technical data from a major supplier (EURECA Messtechnik GmbH, Am Feldgarten 3 D-50769 Köln, GERMANY):

Micro Peltier Elements

Module	I_{\max} [A]	Q_{\max} [W]	U_{\max} [V]	dT_{\max} [K]	Dimensions					Unit Price [Euro]
					A [mm]	B [mm]	C [mm]	D [mm]	H [mm]	
TECM-4-4-1b/69	1,4	0,7	0,9	69	4,3	4,3	4,3	4,3	2,95	28,75
TECM-4-5-1/67	0,7	0,4	1,0	67	3,4	3,4	3,4	5,0	2,30	29,50
TECM-5-7-1/67	0,7	0,9	2,2	67	5,0	5,0	5,0	6,6	2,30	38,50
TECM-7-8-2/67	0,7	1,7	3,9	67	6,6	6,6	6,6	8,3	2,30	52,50
TECM-9-12-4/67	0,7	3,5	8,0	67	9,1	9,9	9,1	11,5	2,30	66,25
TECM-12-6-4/69	1,7	4,4	4,3	69	6,0	12,0	6,0	12,0	2,75	57,50

In this class you will find elements with various geometries and electrical parameters. For this reason, these elements are suited for very different and partly exotic requests as you have in the research. Our support will help you with the selection and the startup of the elements in consideration of your particular requests.

Generally you will receive these elements in small quantities from stock.

Simple Exponential functions

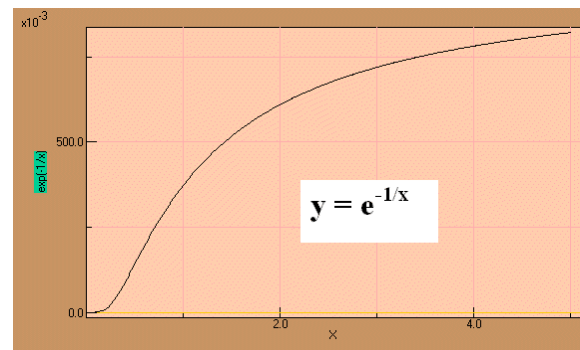
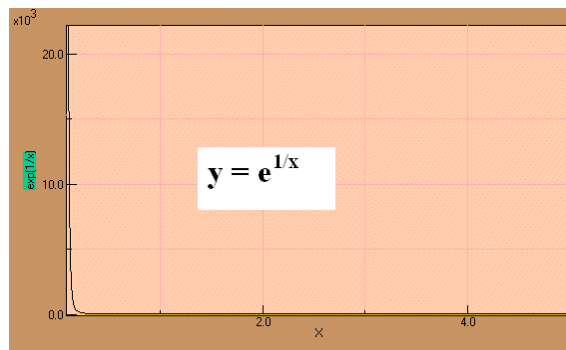
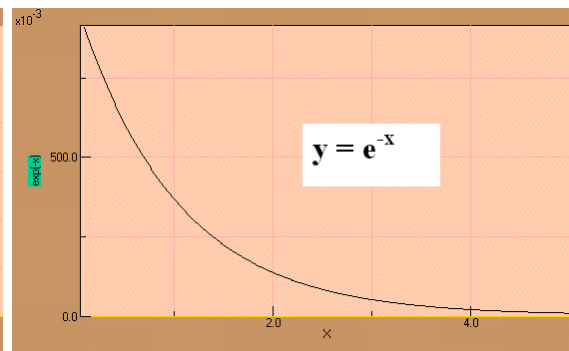
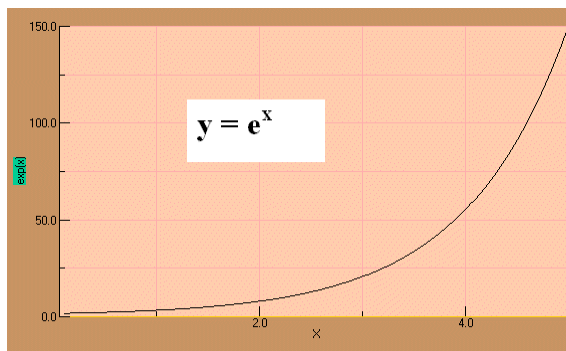
With this JAVA module you can plot exponential (or other) functions

Here are a few examples for writing equations

`cos(x)`
`cos(x*2)`
`cos(x/40)`
`cosh(x)`
`cosh(40*x)`

`exp(x)`
`exp(1/x)`
`x^2`
`1/x^2`
`1/(x^3+10)`
`1/(exp(x-20)+1)`

Here are a few results for exponentials:



Multiple Choice Test zu

2.1.1 Metals

Start Multiple Choice

Multiple Choice Test zu

2.1.2 Alloys

Start Multiple Choice

Multiple Choice Test zu

2.1.3 Non-Metals

Start Multiple Choice

Multiple Choice Test zu

MC 2.1.4 Conductors - Definitions and General Properties

Start Multiple Choice

Multiple Choice Test zu

2.3 Electron Emission

Start Multiple Choice

Multiple Choice Test zu

2.3.3 Thermoelectric Effects

Start Multiple Choice

Multiple Choice Test zu

2.3 Electron Emission and Thermoelectric Effects

Start Multiple Choice

Multiple Choice Test zu

2.4 Ionic Conductors

Start Multiple Choice

Multiple Choice Test zu

MC 2. Conductors

Start Multiple Choice

"Exercise 2.3-1" Quizzes to [Li Ion Battery](#)

Illustration

Quiz 1: How much energy is contained in **1 liter** of gasoline?

Find out approximately by using data you know about your car.

- Consider what you know about your car or some average car: Mileage (let's say **10 km/l**), how far you get on this (**500 km**), engine power (**100 kW**), engine efficiency (??? - think) how long it takes you to go the **500 km** and how much percent of the engine power you use for this on average (you are not always going full throttle!).
- Energy, by the way, is power times time.

Quiz 2: How large or small are **10 kWh /l** - in comparison to known energy (densities).

- How far up do you have to drag your body to gain **10 kWh**? Potential energy, by the way, is $E_{\text{pot}} = mgh$.
- You store energy by pulling up a ball of lead in your (now unused) chimney (diameter = **20 cm**). How far up do you have to pull it to store the **10 kWh**? (Specific density of lead = **11,4 g/cm³**).
- How much water can you bring from **10 °C** to a boil? The specific heat capacity of water is **4.2 kJ/kgK**. A Joule **J**, as we know, is equal to **1 Nm** or **1 Ws**
- Your power output on a bicycle for some length of time is **??? kW**. Mine is at best \approx **100 W**. How far and / or how long do you have to ride your bike before you used up **10 kWh**? How many Big Mac's have you worked off then (a Big Mac of about **120 g** contains **2071,74 kJ = 494,94 kcal**).
- How much energy do you convert into heat if you wrap your car around a tree at **100 km/h**. In other words: $\frac{1}{2}mv^2 = \text{??? kWh}$ (without the gasoline burning). Your car weighs about **1 500 kg** + passengers
- How fast do you need to hit the tree so that the kinetic energy contained in **50 l** \approx **50 kg** of gasoline equal the kinetic energy?

Quiz 3: This and That

- Why do you have plenty of room in an electric car for the battery?
- Why do you want your batteries not to be extremely light weight in a submarine?
- Why is the current collector in a powerful battery not to be sneered at? Hint: Your electric car has a **100 kW** motor; a battery cell has a voltage of about **3 V**.
- You want to be able to run your **100 kW** electric motor car for **5** hours at half the capacity? What is the required battery size? How long does it take you to recharge that battery from your standard outlet (**230 V AC; 16 A**)?
- Take a **1 kWh Li** ion battery with a voltage of **3V**. How much charge **Q** must be stored in it when it is fully loaded?
- How much **Li** atoms would that be? What is their weight? Compare to the [energy density figure](#) and discuss.

Solutions are found in the [Li Ion Battery](#) module

Exercise 2.4.1 Field Screening

Consider a capacitor made of parallel metal plates in air with an area of **1 cm²** at a distance of **1 cm**.

1. How many electrons per **cm²** do you need on the surface to provide for some field ***E*** ending there?
2. What would be the maximum charge density for reasonable field strengths up to an ultimate limit of about **10 MV/cm**? (For higher field strengths you will get violent discharge).
3. How does this number compare to the average *volume* density of electrons in metals. Consider, e.g., from how far away from the surface you have to collect electrons to achieve the required surface density, if you allow the volume density in the afflicted volume to decrease by ***x* %**?

Illustration



Link to the [solution](#)

Solution to Exercise 2.4.1

Illustration

1. How many electrons per cm^2 do you need on the surface a capacitor made of parallel metal plates in air with an area of 1 cm^2 and distance 1 cm to provide for some field E ending there?

The relation between the field E resulting from a homogeneous two-dimensional charge distribution and the charge density ρ_{area} is

$$E = \frac{Q}{\epsilon_0 \cdot A} = \frac{\rho_{\text{area}}}{\epsilon_0}$$

with Q = charge in [C], A = area considered, ρ_{area} = areal density of the charge.

Compare the two formulas for the capacity C of the capacitor formed by the parallel plates if you are not sure about the equation above. We have:

$$C = \frac{\epsilon_0 \cdot A}{d} = \frac{Q}{U} \text{ and } E = \frac{U}{d}$$

2. What would be the maximum density for reasonable field strengths up to an ultimate limit of about 10 MV/cm ? (For higher field strengths you will get violent discharge).

Lets look at some numbers ($\epsilon_0 = 8,854 \cdot 10^{-14} \text{ C/Vcm}$)

Field strength	10^3 V/cm (rather low)	10^5 V/cm (breakdown limit of "normal dielectrics")	10^7 V/cm (close to ultimate breakdown limit)
ρ_{area}	$8,85 \cdot 10^{-11} \text{ C/cm}^2$	$8,85 \cdot 10^{-9} \text{ C/cm}^2$	$8,85 \cdot 10^{-7} \text{ C/cm}^2$

3. How does this number compare to the average volume density of electron in metals. Consider, e.g., from how far away from the surface you have to collect electrons to achieve the required surface density, if you allow the volume density in the afflicted volume to decrease by $x\%$?

The average volume density of electrons in metals is about 1 electron/atomic volume.

Lets keep thing easy and take for the size $d_{\text{atom}} = 1 \text{ \AA}$, which gives $1 \text{ \AA}^3 = 10^{-3} \text{ nm}^3$ for the volume of one atom. The **volume density** of atoms or electrons per cm^3 is thus $\rho_{\text{volume}} = 10^{24} \text{ electrons/cm}^3$.

The **areal density** is whatever is contained in a volume with an extension of just one atom diameter in one direction, i.e.

$$\rho_{\text{areal}} = \rho_{\text{volume}} \cdot d_{\text{atom}} = 10^{-17} \text{ electrons/cm}^2$$

If we want to collect a surplus charge of $Q_{\text{surplus}} = 8,85 \cdot 10^{-7} \text{ C/cm}^2$, the maximum charge from above, from a volume V_{surplus} by reducing the concentration of $10^{24} \text{ electrons/cm}^3$ by $x\%$, we have

$$Q_{\text{surplus}} = \frac{\rho_{\text{volume}} \cdot d}{100 \cdot x}$$

$$d = \frac{100 \cdot x \cdot Q_{\text{surplus}}}{\rho_{\text{volume}}} = \frac{x \cdot 8,85 \cdot 10^{-5}}{10^{24}} \text{ cm} = x \cdot 8,85 \cdot 10^{-29} \text{ cm}$$

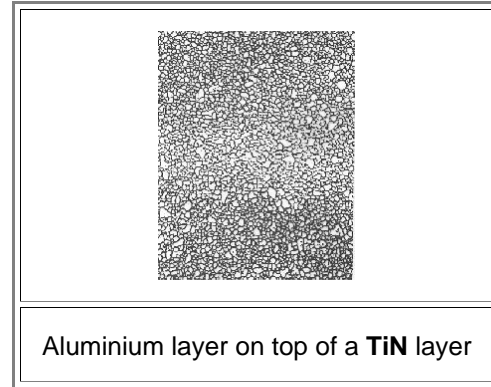
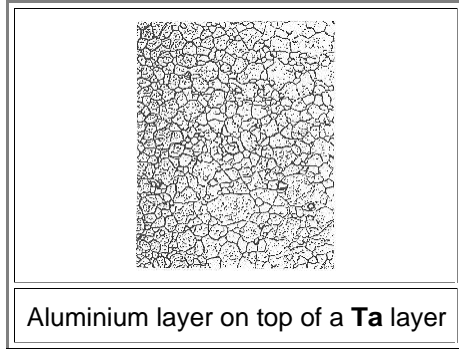
In words: Whatever value we like for x , we only have to change the volume concentration of the electrons in an extremely thin layer a tiny little bit to produce **any** areal charge densities needed - in **metals**, that is!

Al Grain Structure on Different Substrates

Illustration

Around the late eighties, the necessity came up to use a **diffusion barrier** between the **Al** - metallization and the **Si** substrate because the reaction of **Al** with the **Si** in contact holes with cross sections $< 1 \mu\text{m}^2$ became a problem. One material of choice was **TiN**, another one **Ta**.

- The grain structure of the **Al** layer (and with it other properties, e.g. the electromigration resistance, depends significantly on the substrate).
- Below you can see the representative pictures (identical scale) that illustrate this point.



Close examination revealed that the substrate influences:

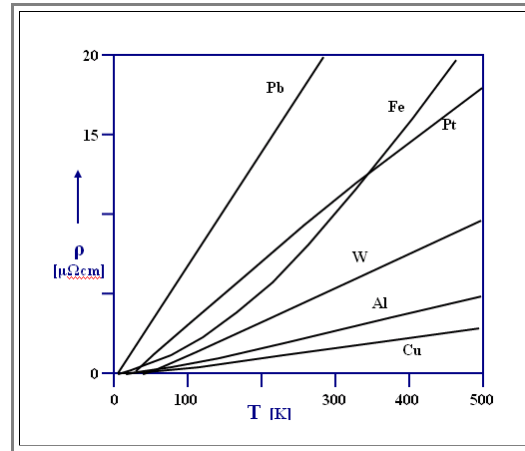
- Grain size
- Grain size distribution
- Texture
- Degree of **Si** precipitation
- Macroscopic stress
- Microscopic stress

All of these properties may influence the performance of the **Al** conductor - and this gives you an idea of what it means to introduce a new material into a fine-tuned product.

Specific Resistivity as a Function of Temperature

Here is a plot of the specific resistivity vs. temperature for some metals.

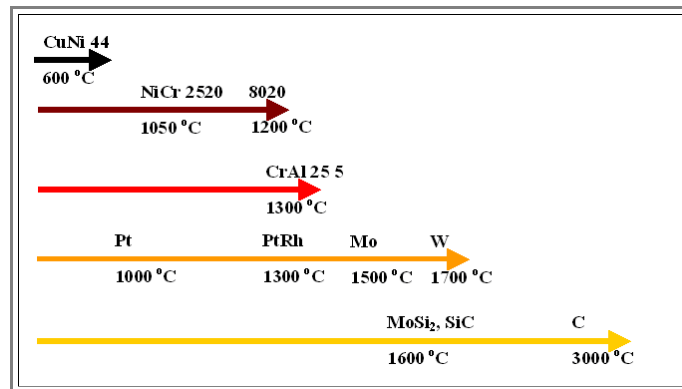
- The resistivity seems to go to **0** for **0 K**. This is not really true, cf. the [detailed curve for Na](#).
- Still, simple free electron theory predicts zero resistance at **0 K**, and trying to see if that is true was one of the driving forces behind the discovery of superconductivity by **Kammerlingh Onnes** in **1911**.



Material for Heating Elements

The figure shows some materials for heating elements and the maximal temperature they can endure.

- Note that the maximum temperature, while more or less scaling with the melting point, is usually well below the melting point - because most materials start to develop all kinds of problems long before they melt.



Illustration

Some Properties of Metals

The following table lists important metal properties (at room temperature).

- Specific resistivity ρ
- Density d
- Temperature coefficient α_ρ of the specific resistivity
- Thermal conductivity λ

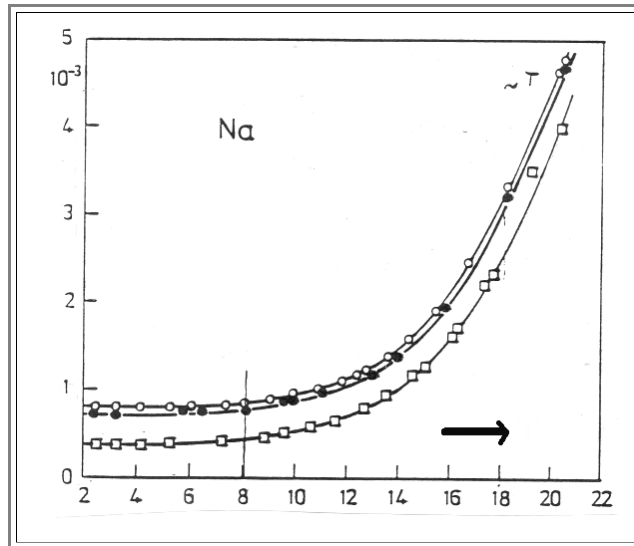
Group	Metal	ρ $\mu\Omega\text{cm}$	d g/cm^3	α_ρ $\%/^\circ\text{C}$	λ $\text{W/cm}^\circ\text{C}$
Ia	Na	4,2	0,97		1,4
	K	6,2	0,86		0,9
Ib	Cu	1,7	8,9	0,43	4,0
	Ag	1,6	10,5	0,41	4,1
	Au	2,2	19,3	0,40	3,1
IIa	Mg	4,5	1,7	0,41	1,4
	Ca	3,9	1,5	0,42	
IIb	Zn	5,9	7,2	0,42	1,1
	Cd	6,8	8,6	0,42	1,0
	Hg	97	13,5	0,08	0,08
IIIa	Al	2,7	2,7	0,43	2,3
IVa	Sn	12	7,3	0,43	0,7
	Pb	21	11,3	0,35	0,4
VIIIb	Fe	9,7	7,9	0,65	0,7
	Co	6,2	8,9	0,60	0,7
	Ni	6,8	8,9	0,69	0,9
Vb/VIb	Ta	13	16,6	0,38	0,5
	Cr	14	7,2	0,30	0,7
	Mo	5,2	10,2	0,40	1,4
	W	5,5	19,3	0,40	1,6
VIIIb	Rh	4,5	12,5	0,42	0,9
	Pd	9,8	12,0	0,38	0,7
	Pt	9,8	21,4	0,39	0,7

Residual Resistivity and Defects


Shown here are measured curves of the low temperature resistivity of **Na** with different defect concentrations

- The resistivity is constant for very small temperatures.
- In the "bend" it shows T^5 characteristics.
- For most of the temperature range it is proportional to T .

Defects clearly do increase the residual resistance (the upper two curves are for **Na** with defects, the lower one for rather perfect **Na**); the effect can be much large in other metals or for larger defect densities.



Some Properties of Cu Alloys

 The following table lists just a few of the existing **Cu** based alloys for specific uses.

Illustration

Alloy	Essential Advantage	σ 10^{-4} S/cm^1	Uses
Cu + 0,2% Ag	relatively hard	57	Collector lamellae
Cu + 0,5% Te	easy machinable	55	precision parts
Cu + 1,2% Cd	high strength	52	Overhead wires for trains
Cu + 0,6% Cr	can be hardened	48	Welding electrodes
Cu + 1,7% Be	can be hardened	18	Contact springs

¹⁾ **S = Siemens = $1/\Omega$**

Some Properties of Ag Based Contact Materials

The following table lists just a few **Ag** based alloys that are obtained by a *sintering process* for specific contact uses

The melting point is always around **960 °C**

Alloy	ρ [$\mu\Omega\text{cm}$]	Brinell Hardness [kp/mm ²]	Uses
Ag (60 ... 90) % Ni (40 ... 10) %	2,8...2,0	50 ... 130	Low Voltage circuit breakers
Ag 88% CdO 12%	2,4	60 ... 70	Circuit breakers
Ag 95% SnO ₂ 5%	2,5	50 ..60	High load relays
Ag 98 % C 2 %	2,5	35 ... 45	Welding resistant contacts
Ag (10 ... 30) % W (90 ... 70) %	5,0 ... 4,0	105 ... 230	Low voltage high power switches

The next table lists just a few **Ag** based alloys that are obtained by *melting* the constituents

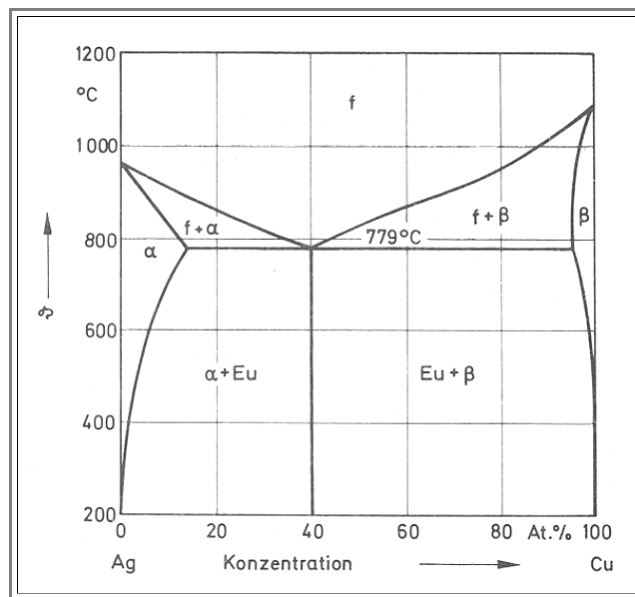
The melting point varies from **(800 - 1030) °C**

Alloy	$\rho\mu\Omega\text{cm}$	Brinell Hardness (kp/mm ²)	Uses
Ag 85 % Cd 15 %	4,8	40...78	Welding resistant DC contacts
Ag (72 ...97) % Cu (28...3) %	1,8 ... 2,1	40...140	Heavy duty relay contacts
Ag 95% Ni 5%	1,9	50	Contacts in communication technology
Ag (70..50) % Pd (30...50) %	15,6...32	70...190	Contacts in communication technology
Ag 20 % Au 80 %	10	37...90	Fine contacts
Cu ca. 95 % Co (...2,5) % Be (1..3) %	3,5...10	100...400	Brushes, pantographs

¹⁾ S = Siemens = 1/ Ω

Phase Diagram Ag - Cu

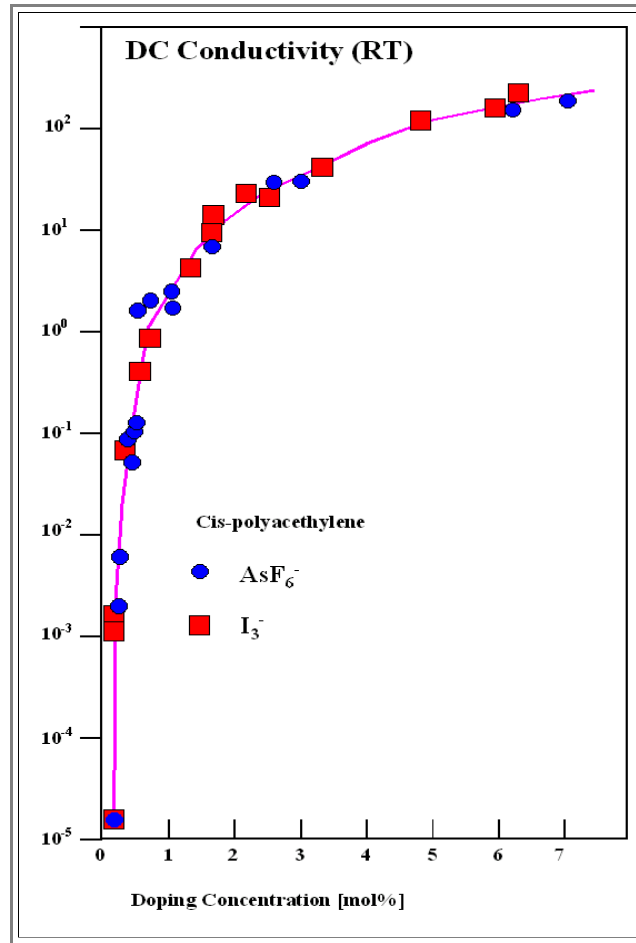
For illustration purposes the **Cu - Ag** phase diagram is shown, demonstrating a miscibility gap.



Conducting Polymers

Shown below is the conductivity of a polymer (cis-polyacetylene) doped with AsF_6^- or I^-

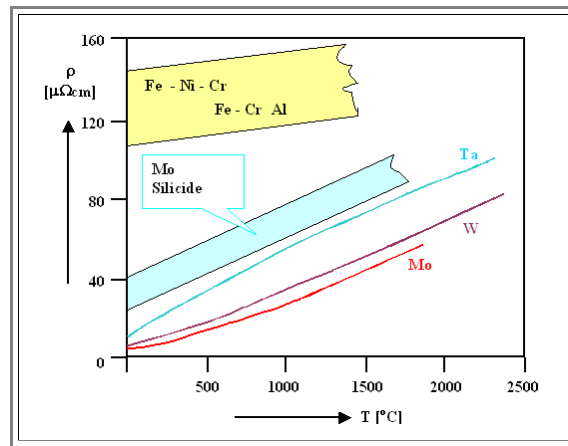
About 1% of doping materials incorporated into the polymer increases the conductivity by 5 orders of magnitude!



Overview of Heating Materials

The diagram shows the temperature range and the resistivity of typical materials for heating elements.

- Do not forget that humble **1500 °C** (roughly the melting point of iron) means that it is *really* hot! Mankind had to invent pretty tricky crafts for about **2000** years because nobody could reach the melting point of iron.
- Beyond **2000 °C** there are not many usable materials left. Essentially you are stuck with graphite (not in the figure).



Example for MoSi₂ Heater

Here the "News" concerning **MoSi₂** heating elements from a German company im **2004**.



Illustration

Elektrische Heizelemente aus MoSi₂ für Industrie- und Laboröfen

06.07.04: Neu in der MolyCom-Serie sind die MolyTec-1600-Heizpaneele, elektrische, modular aufgebaute Beheizungssysteme bis **1600 Grad Celsius**. Einsatzfeld der hoch effizienten MolyTec-Elemente sind Schmelzwannen und Feeder in der Glasindustrie und anderen Hochtemperaturbranchen. MolyTec-Paneele bestehen aus MolyCom-Heizelementen, die in hochtemperaturfesten UltraBoard-Mullitfaserplatten eingebettet sind. UltraSiC-Rohre verstärken die Paneele für eine sichere Aufhängung im Zielsystem.

Die M. E.-SCHUPP-Marke MolyCom fasst elektrische Heizelemente mit einem sehr günstigen PreisLeistungsverhältnis zusammen. Ihr Einsatzbereich reicht bis **1800 Grad Celsius**. Sie setzen sich aus Molybdändisilizid und einigen Keramik-Komponenten zusammen. Über 900 Grad Celsius bildet sich auf ihrer Heizleiteroberfläche eine dichte, aber dünne Quarzglas-Schicht. Sie schützt das MoSi₂- Glühteil vor Oxidation. MolyCom hat sich als technisch gleichwertige, aber kostengünstigere Alternative zu den marktführenden Heizelement-Produkten auf MoSi₂-Basis etabliert.

Durch die Hot-Bending-Technologie (MolyCom-Ultra) können im neuen Werk des Produktionspartners von M. E. SCHUPP nun auch komplexere Heizelemente hergestellt werden, zum Beispiel flache oder zylindrische Multishank-Module, wie sie in den MolyTec-Heizpaneelen zum Einsatz kommen. Ebenfalls möglich sind jetzt größere gebogene Heizelemente mit 90 Grad-Winkeln mit Standarddurchmessern von 3/6 mm bis 12/24 mm.

usw.

Strange words

▀ The somewhat unusual word "*ken*" is best remembered by a variant of a famous English poem:

The Star or **Twinkle, Twinkle, Little Star**
by *Jane Taylor*, 1806

Advanced

Original:

Twinkle, twinkle, little star,
How I wonder
what you are.
Up above the world so high,
Like a diamond in the sky.







Physicists Version:

Twinkle, twinkle little star.
I don't wonder
what you are,
'cause by spectroscopic *ken*
I know you are hydrogen.

▀ *Ken* (Scottish) To know, to recognize (German "Kennen"); To have an understanding of something.







Paul Langevin

Advanced

-  Paul Langevin (1872 – 1946) was one of the fathers of magnetism.
 -  In the "miracle year" 1905, when Einstein published three revolutionary papers, Langevin published his theory of para- and diamagnetism.
 -  Considering that quantum theory had yet to be discovered, and that statistical thermodynamics was in its infancy, his contribution was of amazing insight.
 -  As far as paramagnetism was concerned, he first established the treatment we used in the text for orientatio polarization (which is formally identical to paramagnetism).
-  The resulting function of the form $L(a) = -1/a + \coth(a)$ is therefore called the Langevin function (or sometime Langevin-Debye function, since Debye used the same procedure somewhat later for polar molecules).
-  Langevin also worked in the field of Brownian motion (like Einstein). His work concerning ultrasonic sound during world war I was the base for sonar technology .

Beer and Relaxation

Advanced

-  Around **1970** I was an auxiliary teacher at my old High School, earning some money needed to see me through my study of physics.
 -  Just before christmas, I always conducted a lot of "fun" experiments, including an experiment the students did (I provided the material): Measure the decay of the froth on top of a mug of beer.
 -  Everbody always got the exponential decay law, learned a bit about relaxation phenomena, and just had plain old fun.
-  Little did I know then that this could have gotten me the [\(Ig\)-Nobelprize](#)!
 -  Arnd **Leike** of the University of Munich, won the **2002** physics Ig Nobel prize for demonstrating that beer froth obeys the mathematical Law of Exponential Decay. [REFERENCE: "Demonstration of the Exponential Decay Law Using Beer Froth," Arnd Leike, European Journal of Physics, vol. 23, January 2002, pp. 21-26.]
INTERDISCIPLINARY
 -  Here is a link to this [groundbreaking paper](#), showing that relaxation with beer may have more meanings than you would have imagined.

Frequency Dependence of Orientation Polarization

Advanced

How do we get to the time- and frequency dependence $P(t)$ and $P(\omega)$, respectively, of the orientation polarization without "cutting corners" as in the backbone?

- While in principle each function is just the Fourier transform of the other, it is not so easy to actually do the required math. It is probably best, to start with the differential equation that describes the system.

Within the "relaxation time approximation" always used for those cases we have

$$\frac{dP(t)}{dt} = -\frac{P}{\tau} + S(t)$$

- $S(t)$ is some disturbance or signal or input - whichever term you prefer - that has some time dependence. We need it because otherwise the system would be "dead" and not doing anything after at most one decay if we pick suitable starting conditions.

Whatever happens, we can always write $P(t)$ and $S(t)$ as a [Fourier series, or, more general, as Fourier integral](#) of the correlated $P(\omega)$ and $S(\omega)$ "spectra" of the time functions. Doing this we have

$$P(t) = \int_{-\infty}^{\infty} P(\omega) \cdot \exp(i\omega t) \cdot d\omega$$

$$S(t) = \int_{-\infty}^{\infty} S(\omega) \cdot \exp(i\omega t) \cdot d\omega$$

- $P(\omega) \cdot \exp(i\omega t)$ and $S(\omega) \cdot \exp(i\omega t)$ are the Fourier "coefficients" (with values for every ω , not just harmonics as in Fourier series) for the time functions.

We now have a linear differential equation that is solved by some $P(t)$ which can be expressed as a Fourier transform and that implies that all Fourier coefficients (and any superposition thereof) also solve the differential equation. Inserting the Fourier coefficients directly gives

$$\frac{d\{P(\omega) \cdot \exp(i\omega t)\}}{dt} = i\omega \cdot P(\omega) \exp(i\omega t) = -\frac{P(\omega) \cdot \exp(i\omega t)}{\tau} + S(\omega) \cdot \exp(i\omega t)$$

- From this we obtain

$$P(\omega) \cdot (i\omega + 1/\tau) = S(\omega)$$

If we define $\omega_0 = 1/\tau$ (or $= A/\tau$ if we want to be more general) we now have a simple relation between the Fourier components of input S and output P :

$$P(\omega) = \frac{1}{\omega_0 - i\omega} \cdot S(\omega)$$

The dielectric function $\epsilon(\omega)$ that we are trying to calculate is simple the relation between the output $P(\omega)$ and the input $S(\omega)$, what we get is

$$\epsilon(\omega) := \frac{P(\omega)}{S(\omega)} = \frac{1}{\omega_0 - i\omega}$$

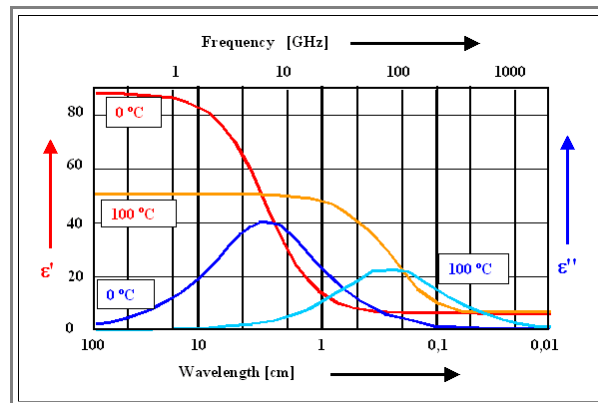
- Considering that the disturbance $\mathbf{S}(\omega)$ must have the dimension of an electrical field, forces us to conclude that it actually must be an electrical field, and we could just as well write $\mathbf{E}(\omega)$. What we have then is essentially the dielectric function as [discussed in the backbone](#).

Cooking with Microwaves

Advanced

When you heat up food with a **microwave oven**, all you do is to jiggle the water molecules with a frequency where the **dielectric losses** are large.

Lets look at the **dielectric function of water**:



First, we see that the general graph of the frequency dependence is exactly as we would expect from the [theory](#).

There is a noticeable change of ϵ' and ϵ'' with the temperature. This is as it should be, we had a temperature dependence for the [polarization](#).

$$\langle P \rangle = \frac{N \cdot \mu^2 \cdot E}{3kT}$$

Changing the temperature from about **300 K** to **400 K** thus should reduce '**(300 K) \approx 80** by about $\frac{3}{4}$; i.e. to **$\epsilon'(400 \text{ K}) = 60$** . The observed reduction is somewhat more severe, because we did not take into account that the water dipoles interact to some extent; and at low temperature this interaction is stronger than at high temperatures.

The maximum of the dielectric losses occurs roughly between **5 GHz - 100 GHz.**, i.e. in the microwave region of the spectrum

Most microwave ovens work at **2.45 GHz**, well below the region of maximum losses. This is intentional to ensure that the radiation is not totally absorbed by the first layer of water it encounters but may penetrate further into the foodstuff, heating it more evenly.

Radiation passing through the food items is mostly reflected back, due to the design of the microwave oven, and absorbed on later passes.

If the water is frozen, i.e. you have ice, you have problems. The **DK** of ice at the microwave frequency is only about **3**, and little energy is absorbed

If your water is salty, the **DK** goes down, too. On the other hand, the **Na⁺** and **Cl⁻** ions are jiggled by the electrical field; too (without producing dipoles and thus a **DK**), producing frictional heat and increasing the heating rate.

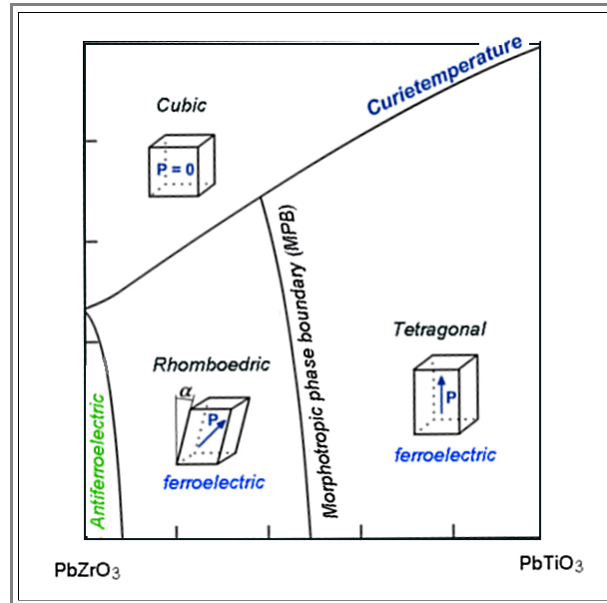
Thus whereas ice is a very poor microwave absorber, and water becomes a poorer microwave absorber with rising temperature, a lossy salty food such as salt meat becomes a better microwave absorber with rising temperature.

This is particularly noticed on thawing. It may take forever to thaw your frozen butter or steak, but after it happened it only takes a short time to turn it liquid or mushy, respectively

PZT or Lead-Zirconate-Titanate

PZT is short for Lead (= **P**b) zirconate (= **Z**r) titanate (= **Ti**) or $\text{Pb}[\text{Zr}_x\text{Ti}_{1-x}]\text{O}_3$. It is mixture of PbTiO_3 and ZrTiO_3

- Both constituents are Perovskites as described in the [backbone](#). For temperatures below the Curie temperature we have spontaneous polarization as indicated in the phase diagram (adopted from Wikipedia).



Both lattice pictures in the phase diagram are hugely exaggerated. Drawn to scale the eye would not see a difference to a perfect cube. The relation of the axis' in the tetragonal case is about **1,06** and the angle α in the rhomboedric structure, for example, is around **0.3°**.

- This rather small distortions are large enough, however to produce major permanent polarization effects

The interesting part is the "**morphological phase boundary**" (**MPB**) where the structure changes but not the constituents. Several properties, most interesting for us the "dielectric constant" show pronounced maxima on the **MPB**; "theoretically" they could diverge to infinity.

- More down to earth, a mixture of tetragonal and rhomboedric crystals have all together **14** possible directions for spontaneous polarization. In other words, no matter what the external field direction might be, there is always an "easy" direction available in the **PZT** that is not too steeply inclined to the field direction.

As always, optimizing the mixture and adding some other "dopants" or better alloying elements, can produce a large range of properties.

Gauss Law or Integral Theorem

Basics

Gauss law relates the **charge** contained inside a volume **V** bounded by a surface **S** to the **flux of the electrical field**.

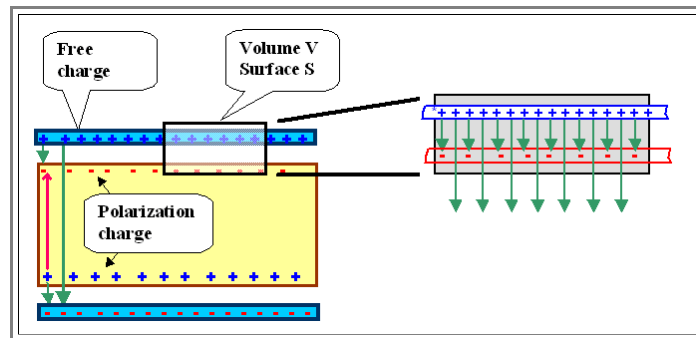
- The **flux of the electrical field** through a surface **S** is the integral over the components of **E** perpendicular to the surface.
- The most simple way to visualize this is to equate the flux with the number of field lines running through the surface.
- The charge is usually expressed in terms of charge density $\rho(x,y,z)$.

Gauss law then states:

$$\iint_S \mathbf{E} \cdot \mathbf{n} \cdot d\mathbf{a} = \frac{1}{\epsilon_0} \cdot \iiint_V \rho(x,y,z) \cdot dV$$

- With \mathbf{n} = normal vector of the surface **S**, $d\mathbf{a}$ = surface increment, dV = volume increment.

Lets apply Gauss law to a capacitor with or without a dielectric inside. We have the following situation:



- Without** a dielectric, **all** green field lines starting at the positive charges of the capacitor plates would run through the interior of the capacitor (and thus through the **lower** surface of the probing volume for applying Gauss' law).
- With** a dielectric inside, only the **"long"** field lines from all field lines starting at the positive charges on the upper electrode will contribute to the **flux of E** because some of the green ones will end at the charges on the surface of the dielectric as shown in the enlargement of the probing volume.

The number of green field lines ending at the surface charge of the dielectric is identical to the number of field lines that we would have inside the dielectric for the given polarization - where green and red ones meet, they cancel each other.

- We see that only the lower surface of our probing volume carries field lines, so the **flux** on this surface is **number of field lines = field times one major area (= A) of the volume V**.
- Without** the dielectric, the flux would be larger because **all** flux lines starting at a positive charge would then contribute. The flux **D** in this simple case would be ,

$$D = E_{ex} \cdot A$$

With **A** = that part of the surface **S** that contains field lines and **E_{ex}** = field caused by the external charges only.

- With** the dielectric, the flux is smaller as reasoned above. We conclude, using Gauss law, that the **amount of charge inside the volume V** must be reduced by the dielectric, which is quite obvious when we look at the picture.

For a quantitative description lets compare the case **with** and **without** dielectric, realizing that the integrations called for in the formulations of Gauss law as given [above](#) are now simple multiplications:

Without dielectric	With dielectric
Electrical flux with Gauss law	

$$D_0 = \epsilon_0 \cdot A \cdot \frac{U}{d} = \frac{Q_0}{\epsilon_0}$$

Gauss law

$$D_{\text{di}} = E \cdot A = \frac{Q_0 + Q_{\text{pol}}}{\epsilon_0}$$

Gauss law

with E_0 = Field without the dielectric,
 U = voltage applied to the capacitor,
 d = distance between plates;
 Q_0 = charge on the plate within V ,
 A = area of the relevant side of V .
 (Only one surface of V contributes)

with E = field inside the capacitor.

$$\frac{Q_{\text{pol}}}{A} \text{ is the polarization } P \text{ by definition.}$$

Electrical flux with Maxwell Definition

Rewriting the equations gives

$$\frac{Q_0}{A} = \epsilon_0 \cdot E_0 \quad := \quad D_0$$

Maxwell definition

$$\begin{aligned} \frac{Q_0}{A} + \frac{Q_{\text{pol}}}{A} &= D_{\text{di}} \\ &= \frac{Q_0}{A} + P \\ &= D_0 + P \\ &:= \epsilon_0 \cdot \epsilon_r \cdot E \end{aligned}$$

Maxwell definition

This is the *definition* of D , the electrical flux density

This is the *definition* of ϵ_r , the **(relative) dielectric constant**

Capacitance

The capacitance C is defined as $C = Q/U = Q/E \cdot d$. Using the equations from above we have

$$C = \frac{Q_0}{E_0 \cdot d} = \frac{A \cdot \epsilon_0 \cdot E_0}{E_0 \cdot d} = \frac{A \cdot \epsilon_0}{d}$$

$$C = \frac{Q}{E \cdot d} = \frac{A \cdot \epsilon_0 \cdot \epsilon_r \cdot E}{E \cdot d} = \frac{A \cdot \epsilon_0 \cdot \epsilon_r}{d}$$

This is of course exactly what we would have expected

Linking the two systems


With $P = \epsilon_0 \cdot \chi \cdot E_0$ it follows

$$\begin{aligned}
 D_{di} &= \epsilon_0 \cdot E_0 + \epsilon_0 \cdot \chi \cdot E_0 \\
 &= \epsilon_0 \cdot (1 + \chi) \cdot E_0 \\
 &:= \epsilon_0 \cdot \epsilon_r \cdot E_0 \\
 \epsilon_r &:= 1 + \chi
 \end{aligned}$$

We thus obtain two simple equations connecting the old-fashioned " D and ϵ_r " world with the modern " P and χ " world

$$D_{di} = D_0 + P$$

$$\epsilon_r = 1 + \chi$$

 **OK** - so it is tedious and boring. And the result is simple and we all know it.

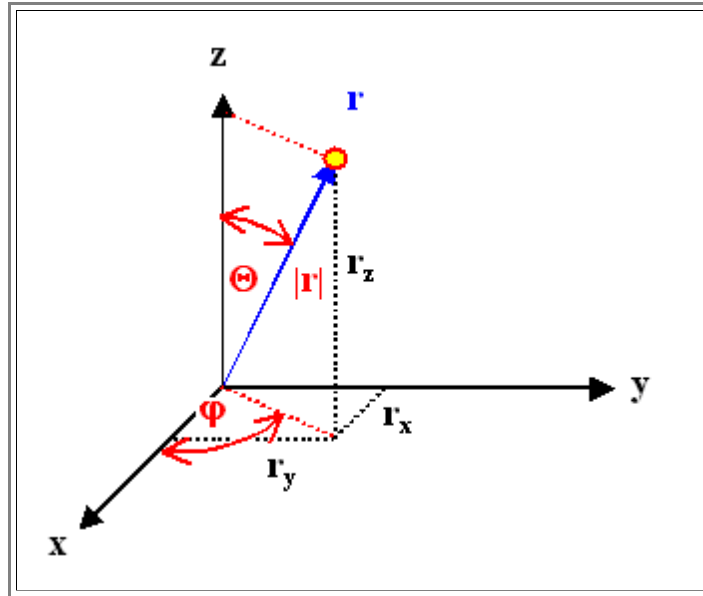
 Still, try at least once in your life to understand *completely* the reasoning behind this. Similar stuff will come up all the time!

Spherical Coordinates

Basics

For many mathematical problems, it is far easier to use spherical coordinates instead of Cartesian ones.

- In essence, a vector \mathbf{r} (*we drop the underlining here*) with the Cartesian coordinates (x, y, z) is expressed in spherical coordinates by giving its distance from the origin (assumed to be identical for both systems) $|\mathbf{r}|$, and the two angles φ and Θ between the direction of \mathbf{r} and the x - and z -axis of the Cartesian system.
- This sounds more complicated than it actually is: φ and Θ are nothing but the geographic **longitude** and **latitude**. The picture below illustrates this.



This is simple enough, for the translation from one system to the other one we have the equations

$$\begin{aligned}
 x &= r \cdot \sin\Theta \cdot \cos\varphi & r &= (x^2 + y^2 + z^2)^{1/2} \\
 y &= r \cdot \sin\Theta \cdot \sin\varphi & \varphi &= \arctg(y/x) \\
 z &= r \cdot \cos\Theta & \Theta &= \arctg \frac{(x^2 + y^2 + z^2)^{1/2}}{z}
 \end{aligned}$$

Not particularly difficult, but not so easy either.

Note that there is now a certain ambiguity: You describe the **same** vector for an ∞ set of values for Θ and φ , because you always can add $n \cdot 2\pi$ ($n = 1, 2, 3, \dots$) to any of the two angles and obtain the same result.

- This has a first consequence if you do an integration. Lets look at the ubiquitous case of normalizing a wave function $\psi(x, y, z)$ by demanding that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(x, y, z) \cdot dx dy dz = 1$$

In spherical coordinates, we have

$$\int_0^{\infty} \int_0^{2\pi} \int_0^{\pi} \psi(r, \varphi, \Theta) \cdot dr d\varphi d\Theta = 1$$

- You no longer integrate from $-\infty$ to ∞ with respect to the angles, but from 0 to 2π for φ and from 0 to π for Θ because this covers all of space. *Notice the different upper bounds!*

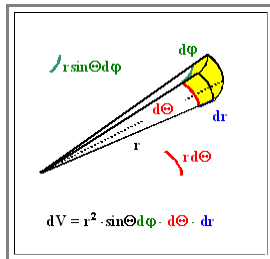
Lets try this by computing the volume V_R of a sphere with radius R . This is always done by summing over all the differential volume elements dV inside the body defined by some equation

- In Cartesian coordinates we have for the volume element $dV = dx dy dz$, and for the integral:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ??? \, dx dy dz$$

- Well, if you can just *formulate* the integral, let alone solving it, you are already doing well!

In spherical coordinates we first have to define the volume element. This is relatively easily done by looking at a drawing of it:



- An incremental increase in the three coordinates by dr , $d\phi$, and $d\theta$ produces the volume element dV which is close enough to a rectangular body to render its volume as the product of the length of the three sides.
- Looking at the basic geometry, the length of the three sides are identified as dr , $r \cdot d\theta$, and $r \cdot \sin\theta \cdot d\phi$, which gives the volume element

$$dV = r^2 \cdot \sin\theta \cdot dr \cdot d\theta \cdot d\phi$$

The volume of our sphere thus results from the integral

$$V_R = \int_0^{\infty} \int_0^{2\pi} \int_0^{\pi} r^2 \cdot \sin\theta \cdot dr \, d\phi \, d\theta = 2\pi \cdot \int_0^{\pi} \int_0^{\infty} r^2 \cdot \sin\theta \cdot dr \, d\theta = 2\pi \cdot [-\cos\theta]_0^{\pi} \cdot \int_0^{\infty} r^2 \cdot dr$$

$$V_R = 2\pi \cdot [2] \cdot 1/3 R^3 = (4/3) \cdot \pi \cdot R^3 \quad \text{q.e.d.}$$

- Not extremely easy, but no problem either.

Next, consider **differential operators**, like **div**, **rot**, or more general, ∇ and $\nabla^2 (= \Delta)$.

- Lets just look at Δ to see what happens. We have (for some function U)

Cartesian coordinates

$$\Delta = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2}$$

Spherical Coordinates

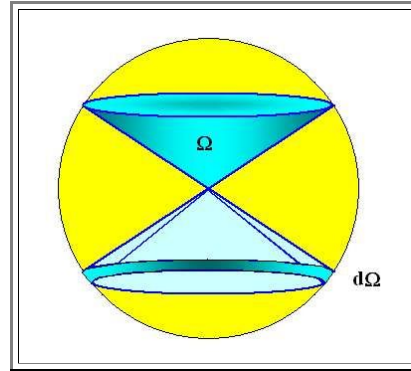
$$\Delta = \frac{\partial^2 U}{\partial r^2} + \frac{2}{r} \cdot \frac{\partial U}{\partial r} + \frac{1}{r^2 \cdot \sin^2 \theta} \cdot \frac{\partial^2 U}{\partial \phi^2} + \frac{1}{r^2} \cdot \frac{\partial^2 U}{\partial \theta^2} + \frac{\cotg \theta}{r^2} \cdot \frac{\partial U}{\partial \theta}$$

- Looks messy, **OK**, but it is still a lot easier to work with this Δ operator than with its Cartesian counterpart for problems with spherical symmetry; witness the [solution of Schrödingers equation for the Hydrogen atom](#).

Looking back now on our treatment of the orientation polarization, we find yet another way of expressing spherical coordinates for problems with particular symmetry:

- We use a **solid angle** Ω and its increment $d\Omega$.
- A **solid angle** Ω is defined as the ratio of the area on a unit sphere that is cut out by a cone with the solid angle Ω to the total surface of a unit sphere ($= 4\pi R^2 = 4\pi$ for $R = 1$).
- A solid angle of 4π therefore is the same as the total sphere, and a solid angle of π is a cone with a (plane) opening angle of 120° (figure that out our yourself).

An incremental change of a solid angle creates a kind of ribbon around the opening of the cone defined by Ω . This is shown below



Relations with spherical symmetry where the value of Θ does not matter - i.e. it does not appear in the relevant equations - are more elegantly expressed with the solid angle Ω .

- That is the reason why practically all text books introduce Θ in the treatment of the polarization orientation. And in order to be compatible with most text books, that was what we did in the main part of the Hyperscript.
- Of course, eventually, we have to replace Θ and $d\Theta$ by the basic variables that describe the problem, and that is only the angle δ in our problem (same thing as the angle φ here).

Expressing $d\Theta$ in terms of δ is easy (compare the [picture in the main text](#))

- The radius of the circle bounded by the $d\Theta$ ribbon is $r \cdot \sin \delta = \sin \delta$ because we have the unit sphere, and its width is simply $d\delta$.
- Its incremental area is thus the relation that we used in the [main part](#).

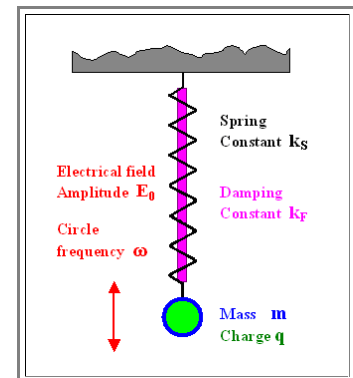
$$d\Theta = 2\pi \cdot \sin \delta \cdot d\delta$$

The Driven, Damped, Harmonic Oscillator

In this module we just recall the essentials of the driven and damped harmonic oscillator - for full details see any textbook of physics, e.g. the [Feynman lectures I-21 - I-25](#).

- We are looking at a system that can be visualized as shown .
- We have a mass m hanging on a spring in the gravitational potential of the earth. We assume in addition that the mass is carrying a fixed charge q .
- The system is **harmonic**, if the force law for the spring is *linear*, i.e. **Hooke's law** applies:

$$F = k_s \cdot x$$

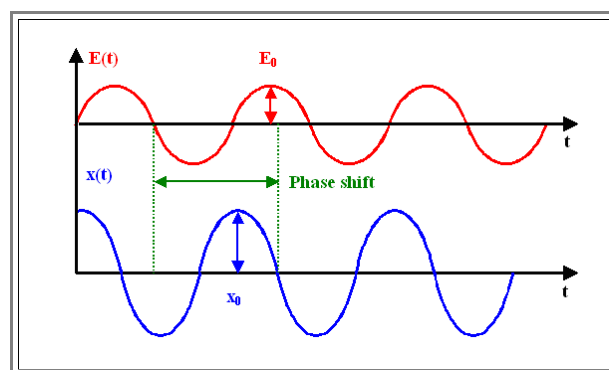


- With F = force acting on the spring, x = elongation caused by F , and k_s = spring constant.
- In other words: The potential U that m "sees" must be parabolic since we always have $F = -dU/dx$.
- We also assume that the system is *damped*, e.g. by a "shock absorber" that is inside the spring like in the suspension system of your car. This is described by damping constant k_f ("F" signifies "friction") and makes sure that an oscillation, once started, will not go on forever.

Finally, we *drive* the system, i.e. we apply a periodically changing driving force - in this example by an oscillating electrical field characterized by its amplitude E_0 and its (circle) frequency ω .

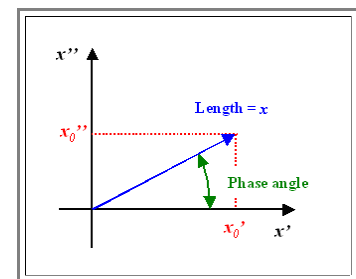
We are interested in two things, as shown below:

- The amplitude x_0 of the oscillation as a function of the system parameters, in particular as a function of the frequency of the driving force.
- The phase of the oscillation relative to the phase of the driving force.



We can rephrase these questions by describing the amplitude of the oscillations as a sum of two sinus function, one exactly in phase with the driving force, and the other one shifted by 90° (which simply makes it a **cosine** function relative to the in-phase **sin** function)

In this case we ask for the in-phase amplitude x_0' and the out-of-phase amplitude x_0'' . The total amplitude x_0 then follows most easily from the "pointer" diagram as shown on the right, we have (with ϕ = phase angle)



$$x = \left(x' + x'' \right)^{1/2} = \frac{x''}{\sin \phi} = \frac{x'}{\cos \phi}$$

By now you realize (it is hoped) that this leads up to the complex notation generally employed for periodic phenomena (otherwise read up on complex numbers and pointer diagrams, or use the [German link](#)).

Anyway, first we have to write down the differential equation for the system. It follows (almost) straight from Newtonian mechanics, we have

$$m \cdot \frac{d^2 x}{dt^2} + k_F \cdot m \cdot \frac{dx}{dt} + k_S \cdot x = q \cdot E_0 \cdot \cos(\omega t)$$

This is simply the force equilibrium and the only non-trivial force in this equation is the term $k_F \cdot m \cdot dx/dt$

This is the damping or friction term, we simply assume that it is proportional to the mass m and its velocity dx/dt . The proportionality constant is our damping constant k_F *times the mass*.

Often the friction term is just written as $k_F^* \cdot dx/dt$, i.e. the mass is included in k_F^* , but our approach has a certain advantage as we will see below.

While all other terms come from ironclad first principle physical law (always assuming harmonic potentials), the friction term is a bit arbitrary; its exact formulation depends on the specific problem.

However, if you have a system where the amplitude "decays" exponentially after the driving force is switched off, you must have a damping term as given. Essentially you are back to the very general model of relaxation into the ground state as employed for the [frequency dependence of the orientation polarization](#).

We are now stuck with solving a linear second order differential equation - and we know how that is done.

Usually, we would move step by step, first looking at a simplified system without damping and driving forces, and then adding the complications.

What we would find for the simplified system is that there is a special frequency ω_0 called the **resonance frequency** or "Eigenfrequency", which is the simply the frequency with which the system will oscillate by itself if started once. The resonance frequency without damping we call ω_0' ; it is given by

$$\omega_0' = \left(\frac{k_S}{m} \right)^{1/2}$$

With damping added, the resonance frequency changes somewhat, and the amplitude will decrease with time after some initial push started an oscillation. This is described by the following equations

$$x(t) = x_0 \cdot \cos(\omega_0 t) \cdot \exp - \frac{k_F}{2} \cdot t$$

$$\omega_0 = \left(\frac{k_S}{m} - \frac{k_F^2}{4} \right)^{1/2}$$

If, for a moment, we apply these equations to an ion sitting in a lattice, we will notice two interesting points:

1. The "spring constant" follows from the binding potential. It is - of course - related to Young's modulus Y which tells us how much the length of a specimen changes under an applied force, or more precisely, how **stress** applied to a material creates (elastic) **strain**. For a homogeneous isotropic material [we actually have](#)

$$k_S = Y \cdot a_0$$

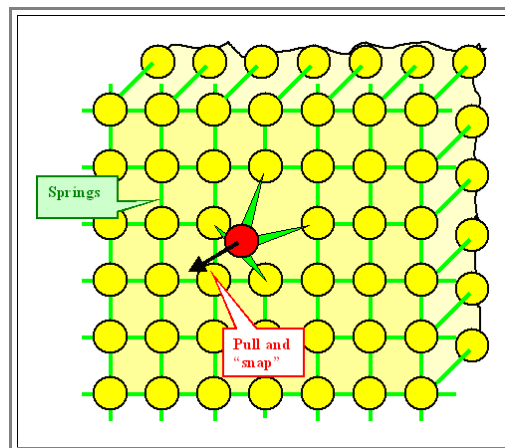
With a_0 = bond length \approx lattice constant. In other words, we know a lot about the spring constant for the systems we are treating here.

What that means is that we also have a good idea for the order of magnitude of the resonance frequency. [It will come out](#) to be roughly 10^{13} Hz.

2. The damping or friction constant k_F for a single atom, which is coupled by "bond springs" to some other atoms, which are coupled by bond springs ... and so on, is far more difficult to assess. Off hand, most of us probably do not have the faintest idea about a possible numerical value, or if k_F relates somehow to some quantities we already know, like the spring constant.

However, realizing that the dimension of the damping constant is $[k_F] = 1/s$, and that it takes just a few reciprocal k_F 's before the oscillation dies out, we can make an **educated guess**:

If you "snap" just **one** atom of a huge collection of more or less identical atoms, all connected by more or less identical springs, pretty soon all atoms will oscillate. And the original energy, initially contained in the amplitude of the "snapped" atom, is now spread out over all atoms - which means that their amplitudes will be far smaller than the original one. To get the idea, just look at the picture.



In other words: There is no doubt that it will just take a few - say **5** or maybe **50** - oscillations of the primary atom, before the **orderly** energy contained in the oscillation of that one atom will have spread and become **diluted** and **disordered**.

In yet other words: excess energy contained in the oscillations of one atom will turn into thermal energy (= random vibrations of **all** the atoms); it becomes **thermalized** rather quickly - in the time it takes to oscillate back and forth a few times.

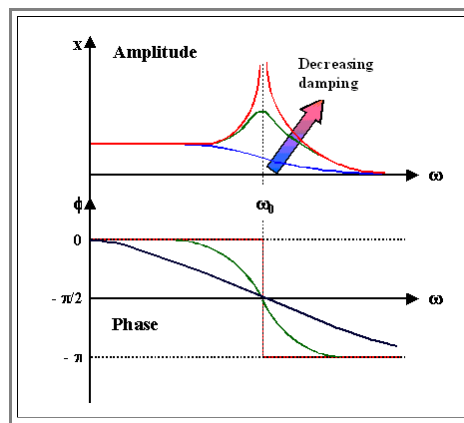
$k_{F\omega}$ is thus tied to ω_0 , we expect it to be very roughly in the order of **$5\omega_0$ $50\omega_0$** .

So far so good. But now we must go all the way and switch on "driving", in our example an electrical field that pulls at the charged mass with a force that oscillates with some arbitrary frequency ω

However, we will not even try to write down the solution the full differential equation given above in "straight" terms - it is too complicated, and there is a better way. We will, however, consider the solution **qualitatively**.

We (should) know that the mass oscillates with the frequency of the driving force and an amplitude that depends on the frequency (and the damping constant and so on), and that there will be a phase shift between the driving force and the oscillating mass that also depends on the frequency, and so on.

We also (should) know what all of this looks like - qualitatively. Here it is:



What we are going to do, of course, is to describe the driven damped harmonic oscillator in **complex notation**. The basic equation then is

$$m \cdot \frac{d^2 x}{dt^2} + k_F \cdot m \cdot \frac{dx}{dt} + k_S \cdot x = q \cdot E_0 \cdot \exp(i\omega t)$$

The solutions are most easily obtained for the in-phase amplitude x_0' and the out-of-phase amplitude x_0'' .

The total amplitude x_0 and the phase shift ϕ are contained in these amplitudes. If we want to have them, we simply calculate them as [outlined above](#).

The solution we will obtain is

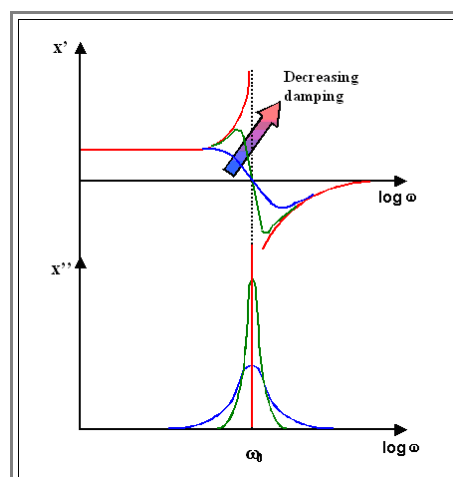
$$x(\omega, t) = x(\omega) \cdot \exp(i\omega t)$$

$$x(\omega) = \frac{q \cdot E_0}{m} \left(\left(\frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + k_F^2 \omega^2} \right) - i \cdot \left(\frac{k_F \omega}{(\omega_0^2 - \omega^2)^2 + k_F^2 \omega^2} \right) \right)$$

$$x'(\omega) = \frac{q \cdot E_0}{m} \left(\frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + k_F^2 \omega^2} \right)$$

$$x''(\omega) = \frac{q \cdot E_0}{m} \left(\frac{k_F \omega}{(\omega_0^2 - \omega^2)^2 + k_F^2 \omega^2} \right)$$

This looks complicated, but is, in fact, far more elegant than the description without complex numbers. If we plot $x'(\omega)$ and $x''(\omega)$, we obtain the following curves



These curves are purely qualitative. A quantitative rendering can be obtained by the JAVA module below

- Instead of the spring constant, you may enter Youngs modulus directly. Typical numbers (in **GPa**) are:
 - Diamond: 1000
 - Carbides, Oxides, Nitrides: \approx 300 - 600
 - Glas: 70
 - Quartz: 100
 - Alkali halides: 15 - 70
 - Wood: 10
 - Polymers: 1 - 10
 - Rubber: 0.001 - 0.1
- The damping constant enters with its reciprocal value normalized to ω , i.e. roughly the number of cycles it takes to dampen out an oscillation.
- You can compare two sets of parameters, because the last curve will always be shown with the new curve.
- You can also enlarge any portion of the diagram by simply drawing a window on the part you like to see enlarged.



Fourier Series and Transforms

Fourier Series

Basics

Every (physically sensible) **periodic** function $f(t) = f(t + T)$ with $T = 1/\nu = 2\pi/\omega$ and ν, ω = frequency and angular frequency, respectively, may be written as a **Fourier series** as follows

$$f(t) \approx a_0/2 + a_1 \cdot \cos \omega t + a_2 \cdot \cos 2 \omega t + \dots + a_n \cdot \cos n \omega t + \dots \\ + b_1 \cdot \sin \omega t + b_2 \cdot \sin 2 \omega t + \dots + b_n \cdot \sin n \omega t + \dots$$

and the Fourier coefficients a_k and b_k (with the index $k = 0, 1, 2, \dots$) are determined by

$$a_k = \frac{2}{T} \cdot \int_0^T f(t) \cdot \cos k \omega t \cdot dt \\ b_k = \frac{2}{T} \cdot \int_0^T f(t) \cdot \sin k \omega t \cdot dt$$

This can be written much more elegantly using [complex numbers](#) and functions as

$$f(t) = \sum_{n=-\infty}^{+\infty} c_n \cdot e^{in \omega t}$$

The coefficients c_n are obtained by

$$c_n = \int_0^T f(t) \cdot e^{-in \omega t} \cdot dt = \begin{cases} \frac{1}{2}(a_n - ib_n) & \text{for } n > 0 \\ \frac{1}{2}(a_{-n} + ib_{-n}) & \text{for } n < 0 \end{cases}$$

The function $f(t)$ is thus expressed as a sum of sin functions with the **harmonic frequencies** or simply **harmonics** $n \cdot \omega$ derived from the fundamental frequency $\omega_0 = 2\pi/T$.

The coefficients c_n define the **spectrum** of the periodic function by giving the amplitudes of the harmonics that the function contains.

Fourier Transforms

A nonperiodic function $f(t)$ ("well-behaved"; we are not looking at some **abominable** functions only mathematicians can think of) can also be written as a Fourier series, but now the Fourier coefficients have some values for all frequencies ω , not just for some harmonic frequencies.


Instead of a spectrum with defined lines at the harmonic frequencies, we now obtain a **spectral density** function $g(\omega)$, defined by the following equations

$$f(t) = \int_{-\infty}^{+\infty} g(\omega) \cdot e^{i\omega t} \cdot d\omega$$

$$g(\omega) = (1/2\pi) \cdot \int_{-\infty}^{+\infty} f(t) \cdot e^{-i\omega t} \cdot dt$$

🔷 The simplicity, symmetry and elegance (not to mention their usefulness) of these **Fourier integrals** is just amazing!

Complex Notation

 Please refer to the [corresponding module](#) (in German) in the Hyperscript "Introduction to Materials Science I" via the link

Basic Optics



Refer to chapter [5. Optics](#) of the Hyperscript "Advanced Materials! You will find:

5.1 Basic Optics

5.1.1 What is Light?

5.1.2 Basic Geometric Optics

5.1.3 Basic Wave Optics

5.1.4 Energy Flow, Poynting Vector and Polarization

5.1.5 Summary to: 5.1 Optics

Fresnel Laws

Refer to chapter **5 Optics** of the Hyperscript "Advanced Materials!"

In subchapter [5.2 Optics and Materials](#) you will find the Fresnel Equations

Multiple Choice Test zu

3.1.1 Dielectrics General

Start Multiple Choice

Multiple Choice Test zu

3.2.1 Polarization Mechanisms

Start Multiple Choice

Multiple Choice Test zu

3.2.2 Electronic Polarization

Start Multiple Choice

Multiple Choice Test zu

3.2.3 Ionic Polarization

Start Multiple Choice

Multiple Choice Test zu

3.2.4 Orientation Polarization

Start Multiple Choice

Multiple Choice Test zu

3.2.5 Dielectrics - Summary and Generalization

Start Multiple Choice

Multiple Choice Test zu

3.2.7 Polarization Mechanisms - Summary

Start Multiple Choice

Multiple Choice Test zu

3.3.4 Complete Frequency Dependence of a Dielectric Material

Start Multiple Choice

Multiple Choice Test zu

3.4.1 Dielectric Losses

Start Multiple Choice

Multiple Choice Test zu

3.5.1 Dielectric Breakdown

Start Multiple Choice

Multiple Choice Test zu

3.6.1 Piezo and Ferroelectricity

Start Multiple Choice

Multiple Choice Test zu

3.7.1 Dielectrics and Optics - Basics

Start Multiple Choice

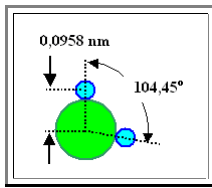
Multiple Choice Test zu

3.1.1 Dielectrics General

Start Multiple Choice

Exercise 3.2-1 Maximum Polarization of Water

Looking into some standard reference book with numbers, e.g. the "*CRC Handbook of Chemistry and Physics*", you find that the structure of a water molecule and its dipole moment is



$$\mu_{\text{water}} = 1,87 \cdot 10^{-18} \text{ e.s.u.}$$

Illustration

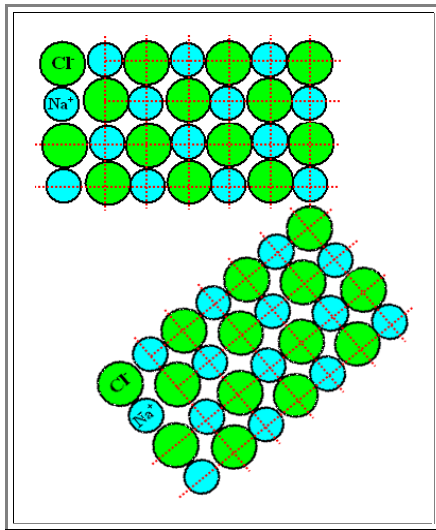
1. How large would the *dielectric constant* of water be if *all* water dipoles are completely oriented into the field direction? Or, if you realize right away that this question doesn't make much sense:
- 1a. How large would the *polarization* of water be if *all* water dipoles are completely oriented into the field direction?
- And what, for gods sake, are **e.s.u.**?

Link to the [solution](#)

Exercise 3.2-2 Interface Polarization

Look at a grain boundary in a simple ionic crystal. Take two grains, for example, that are only rotated with respect to each other and are joined along some grain boundary as shown below

Illustration



Construct a grain boundary by filling in material

Show that in most cases the boundary will be charged and contain dipoles



Link to the [solution](#)

Exercise 3.2-3 Electronic Polarization

Illustration

Look at an atom with atomic number z .

- How large is the distance d between the (center of gravity) of the positive and negative charges for reasonable field strengths and atomic numbers, e.g. the combinations of

- **1 kV/cm**
- **100 kV/cm**
- **10 MV/cm**
- , the last one being about the ultimate limit for the best dielectrics there are,

and

- **$z = 1$** (H, Hydrogen)
 - **$z = 50$** (Sn, (= tin), ...)
 - **$z = 100$** (?)
- Calculate the "spring constant" and from that the resonance frequency of the "electron cloud" (assume the nucleus to be fixed in space).



Link to the [solution](#)

Solution to Exercise 3.2-1

Illustration

Sometimes, a question can be more tricky than originally intended. That is the case here - lets see why.

- First lets get the **e.s.u** out of the way. It means "**electrostatic units**" which are sub-units of the old **c.g.s** (**centimeter-gram-second**) **system**, and still much in use.
- Few things are more confusing than converting electric or magnetic **c.g.s.** units into the **SI (Standard International) kilogram- meter- second-Ampère system**. If you are not somewhat familiar with that, read up the basic modules accessible by the links to this topic.
- In the case given here, you have to multiply with $|c|/10 = 3,3356 \cdot 10^{-10}$ (**c** = vacuum speed of light) to obtain the charge in **[C]** (The magnitude signs **|** simply mean hat you only take the number!); and since the dipole moment is charge times distance, the distance in **e.s.u** units must be **cm**.
- We obtain

$$\mu_{\text{water}} = 1,87 \cdot 10^{-18} \cdot 3,3356 \cdot 10^{-10} \text{ C} \cdot \text{cm} = 6,24 \cdot 10^{-28} \text{ C} \cdot \text{cm}$$
- Lets see if that is reasonable: A water molecule carries about one elementary charge = $1,6 \cdot 10^{-19} \text{ C}$ at the end of the dipole, and the distance will be about $1 \text{ Å} = 10^{-8} \text{ cm}$. This would give a dipole moment of $1,6 \cdot 10^{-27} \text{ C} \cdot \text{cm}$, so the number we got should be correct

Now to the tricky part. First it is important to realize that:

- A material with **completely oriented natural dipoles** does not have a dielectric constant ϵ_r or dielectric susceptibility $\chi = \epsilon_r - 1$ anymore!
- Consider: χ was the proportionality factor between the external field **E** and the induced polarization **P**

$$P = \epsilon \cdot \chi \cdot E$$

- If the field doubles, the polarization, and thus the degree of orientation into the field doubles.
- However, if **all** dipoles are **fully aligned**, the polarization is at a maximum and will not respond to the field anymore; χ loses its meaning.

Nevertheless, we could take this fully polarized material, stick it into a plate capacitor, and just measure how the capacitance **C** changes . This would give us a value for ϵ_r simply by computing **C_{after}/C_{before}**. Lets see if we can do this.

- For the capacity **before** we use our fully polarized dielectric we have with some applied voltage **U** and some corresponding charge **Q₀**

$$C_{\text{before}} = \frac{Q_0}{U}$$

- For the capacity **after** we use our fully polarized dielectric **we have**

$$C_{\text{after}} = (Q_0 + Q_{\text{pol}})/U,$$
 and this gives us

$$\frac{C_{\text{after}}}{C_{\text{before}}} = \epsilon_r = \frac{Q_0 + Q_{\text{pol}}}{Q_0}$$

This does not help, however, because we do not know **Q₀**. Lets try a different approach and look at **C_{after} - C_{before}**.

- We obtain .

$$C_{\text{after}} - C_{\text{before}} = \epsilon_r \cdot C_{\text{before}} - C_{\text{before}} = C_{\text{before}} \cdot (\epsilon_r - 1) = C_{\text{before}} \cdot \chi = \frac{Q_0 + Q_{\text{pol}}}{U} - \frac{Q_0}{U} = \frac{Q_{\text{pol}}}{U}$$

$$\chi = \frac{Q_{\text{pol}}}{U \cdot C_{\text{before}}} = \frac{Q_{\text{pol}}}{Q_0}$$

- This looks better, but it is still not useful - we do not know Q_0 . We still have the same problem: The changes are not *proportional* to what we had *before* the introduction of the dielectric, but *absolute* - we are, in effect, adding a fixed charge and thus switching a second capacitor in series.

Lets try a different approach. We know that $\chi(\text{H}_2\text{O}) \approx 80$. The polarization that goes with this value increases steadily as the field strength inducing the polarization increases - as long as we have $P = \chi \cdot E$

- For large field strength, however, this "law" must break down - we reach the absolute limit of polarization sooner or later.
- So lets compute in a first approximation the field strength needed (within the simple law) to induce the maximum polarization and compare the value obtained to field strengths usually encountered.

First, we compute the maximum polarization P_{\max} . This is simply the the charge $q_{\text{H}_2\text{O}}$ on one end of the water dipole times the distance of the charges $d_{\text{H}_2\text{O}}$ divided by the (area) density of the dipoles, i.e. the (area density) of water.

- The dipole moment of water is given by

$$\mu_{\text{water}} = q_{\text{H}_2\text{O}} \cdot d_{\text{H}_2\text{O}} = 1,87 \cdot 10^{-18} \cdot 3,3356 \cdot 10^{-10} \text{ C} \cdot \text{cm}$$

- We need $d_{\text{H}_2\text{O}}$ to compute $q_{\text{H}_2\text{O}}$; from the picture in the [question](#) we find it to be $d_{\text{H}_2\text{O}} = 0,0958 \text{ nm} \cdot \cos(104,45^\circ/2) = 0,0586 \text{ nm}$.
- The (effective) charge $q_{\text{H}_2\text{O}}$ at the end of a dipole thus is

$$q_{\text{H}_2\text{O}} = \mu_{\text{water}} / d_{\text{H}_2\text{O}} = 6,24 \cdot 10^{-28} \text{ C} \cdot \text{cm} / 0,0586 \cdot 10^{-7} \text{ cm} = 1,065 \cdot 10^{-19} \text{ C}$$

about 2/3 of an elementary charge.

The density of water is $\rho_{\text{H}_2\text{O}} = 1 \text{ kg/l} = 1 \text{ g/cm}^3$ by definition.

- One mol of water is $1 + 1 + 16 = 18 \text{ g}$ which tells us that we have $1 \text{ mol} = 6.022 \cdot 10^{23}$ water molecules in 18 cm^3 .
- The areal density ρ_{areal} of dipoles is therefore

$$\rho_{\text{areal}} = \frac{6.022 \cdot 10^{23} \cdot 0.0586 \text{ nm}}{18 \text{ cm}^3} = 1,96 \cdot 10^{14} \text{ dipoles/cm}^2$$

- Converting *volume* densities to *areal* or *surface densities* may appear tricky. If you are not sure about [how it is done](#), consult the link.
- The maximum polarization P_{\max} thus is.

$$P_{\max} = 1,065 \cdot 10^{-19} \cdot 1,96 \cdot 10^{14} \text{ C/cm}^2 = 2,087 \cdot 10^{-5} \text{ C/cm}^2$$

- If we want to generate this polarization with an electrical field and a susceptibility $\chi = 80$, we need a saturation field strength E_{sat} of

$$E_{\text{sat}} = P_{\max} / \chi \cdot \epsilon_0 = 2,087 \cdot 10^{-5} / 80 \cdot 8,854 \cdot 10^{-12} (\text{C/cm}^2) \cdot (\text{Vm/C}) = 2,946 \cdot 10^6 \text{ V/cm}$$

OK, that is a definite result. Now we have to ask ourselves, how we must compare a field strength of about $3 \cdot 10^6 \text{ V/cm}$ to "normal" field strengths.

- To some extent, we do that in [sub-chapter 3.5.1](#), but common sense tells us that we would certainly use 1 mm or more of a dielectric to insulate a wire carrying 1000 V , for example. This translates to a "typical" field strength of 10.000 V/cm .
- Many materials will be destroyed at field strengths of very roughly 100.000 V/cm , so $3 \cdot 10^6 \text{ V/cm}$ is very large, indeed.
- However, dielectrics in integrated circuits must be able to operate at field strength of this order of magnitude. Take 3 V and a thickness of the dielectric of 10 nm - a not atypical combination - and you have a field strength of $3 \cdot 10^6 \text{ V/cm}$, just what we calculated.

Anyway, if we take 100.000 V/cm as a "normal value, we realize that the only **3,4%** of the dipoles need to be oriented in field direction, whereas the rest could be oriented at random. ($1 \cdot 10^5 / 2,946 \cdot 10^6 = 0,034$).

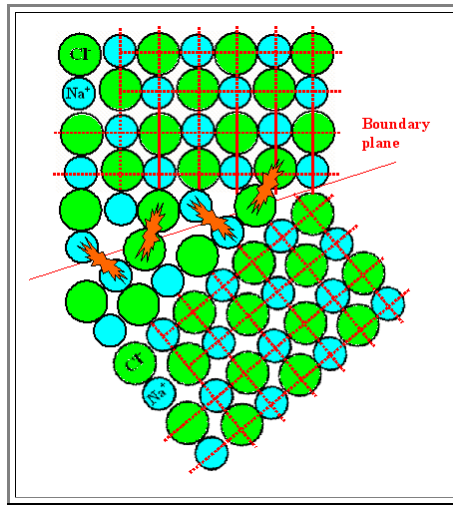
- This is not the physical reality, of course. A more physical interpretation is that all dipoles change whatever orientation they happen to have by about **3,4 %** in field direction. What that means **precisely**, we will leave open, the general meaning, however, is clear:
- The effect of polarization would hardly be noticeable by just looking at the distribution of the dipoles. It is a rather small effect, even for a material with a comparatively very large dielectric susceptibility.

Solution to Exercise 3.2-2

Again, we have a question here that's a bit more tricky than it looks

- Let's look what we get if we construct a grain boundary along the lines suggested.

Illustration



- In this example, some boundary plane was drawn in; then atoms were just added on regular positions to the top to bottom crystal as long as there was enough space.
- The result is not satisfying: We have many "head-on" situations, where atoms with the same charge are in intimate contact. This would be a high-energy situation, and that is unlikely.
- The crystal for almost sure will be smarter than us, and arrange its atoms - ions, to be more precise - in a better way.
- Possibly, also atoms a bit away from the grain boundary need to be re-arranged a bit, for that goal.
- What do we learn from that? Two things:
 1. The precise atomic structure of grain boundaries in ionic crystal (or any crystals with ionic components in their bonding) is not easy to predict. In fact, even the precise structure of grain boundaries in simple metal crystals is far from being simple.
 2. Whatever kind of smart structure the crystal will realize, it is quite likely that there will be some charge imbalance in and around the grain boundary. That means that we also have an imbalance of dipoles - they do not nearly as nicely cancel each other as the neighboring dipoles in the undisturbed lattice.
- And that simply means that grain boundaries - or any kind of interface - will most likely contribute more than the perfect lattice to the polarization of the material. Of course, only for that part of the "volume" that it occupies

Solution to Exercise 3.2-3

Illustration

How large will be the distance **d** between the (center of gravity) of the positive and negative charges for reasonable field strengths and atomic numbers, e.g. the combinations of

- **1 kV/cm**
- **100 kV/cm**
- **10 MV/cm**
- , the last one being about the ultimate limit for the best dielectrics there are,
- **z = 1** (H, Hydrogen)
- **z = 50** (Sn (= tin), ...)
- **z = 100** (?)

From the backbone we have a relation for **d** as a function of **z**, **m** the radius **R** of the atom, and the field strength **E**:

$$d_E = \frac{4 \pi \epsilon_0 \cdot R^3 \cdot E}{ze}$$

We need to look up some number for the radius of the three atoms given (try this link), then the calculation is straight forward - let's make a table:

Atom	R	d(1 kV/cm)	d(100 kV/cm)	d(10 MV/cm)
z = 1				
z = 50				
z = 100				

- Compared to the radius of the atoms, the separation distance is tiny. No wonder, electronic polarization is a small effect *with spherical atoms!*

Calculate the "spring constant" and from that the resonance frequency of the "electron cloud" (assume the nucleus to be fixed in space).

If you don't know off-hand the resonance frequency of a simple harmonic oscillator - that's fine. If you don't know exactly what that is, and where you can look it up - you are in deep trouble.

- Anyway, in [this link](#) you get all you need. In particular the resonance (circle) frequency ω_0 of a harmonic oscillator with the mass **m** and the spring constant **k_S** is given by

$$\omega_0 = \left(\frac{k_S}{m} \right)^{1/2}$$

- How large are the spring constants? That is question already answered in the backbone, so we import the equation

$$k_S = \left(\frac{(ze)^2}{4 \pi \epsilon_0 \cdot R^3} \right)$$

Again, let's make a table for the answers:

Atom	Spring constant	ω_0
z = 1		
z = 50		
z = 100		

Values for dielectric constants

Here are some values for dielectric constants:

- First, some of the more interesting materials with [electronic polarization](#) at work, at least to some noticeable extent.

Illustration

Static ϵ_r for some semiconductors.			
Covalent bonding		Covalent and ionic bonding	
Electronic polarization is the main mechanism		Mix of electronic and ionic polarization	
C	5.7	ZnO	4.6
Si	12.0	ZnS	5.1
Ge	16.0	ZnSe	5.8
SiC	6.7	CdS	5.2
GaP	8.4	CdSe	7.0
GaAs	10.9	BeO	3.0
InP	9.6	MgO	3.0
The numbers are from "Solid State Physics" of Ashcroft / Mermin (an advanced text book).			

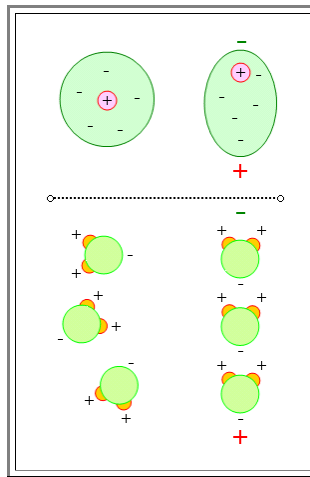
- Next, some numbers for [ionic crystals](#). Besides the static $\epsilon_r(\omega = 0)$, the value $\epsilon_r(\omega = \infty)$ for very high frequencies is also given. This means that after ionic polarization "dies out", there is still some electronic polarization left.

Static and high-frequency ϵ_r for some ionic crystals		
Crystal	$\epsilon_r(\omega = 0)$	$\epsilon_r(\omega = \infty)$
LiF	9.01	1.96
NaF	5.05	1.74
KF	5.46	1.85
LiCl	11.95	2.78
NaCl	5.90	2.34
KCl	4.84	2.19
LiBr	13.25	3.17
NaBr	6.28	2.59
LiI	16.85	3.80
NaI	7.28	2.93
The numbers are from "Solid State Physics" of Ashcroft / Mermin (an advanced text book).		

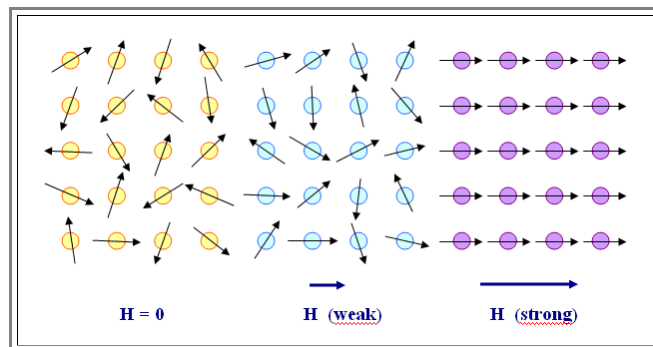
Wrong Illustrations of Orientation Polarization

Illustration

- Some (undergraduate) text books like the "Atkins", "Barrett" or "Askeland" show simplified pictures to orientation polarization that are simply *wrong*.



- This figure reproduces roughly a picture from the (otherwise very good) "[Askeland: The Science and Engineering of Materials](#)". In the bottom part it is supposed to show water molecules without (left) and with (right) an electrical field.
- Well - it is simply wrong. However, considering that this book only spends two pages on polarization mechanisms and dielectric constants, there simply is no room for the author to describe "details".
- The same is true for the magnetic equivalent of orientation polarization: [paramagnetism](#).
- Here is a picture drawn after an illustration in the "[Barrett](#)"

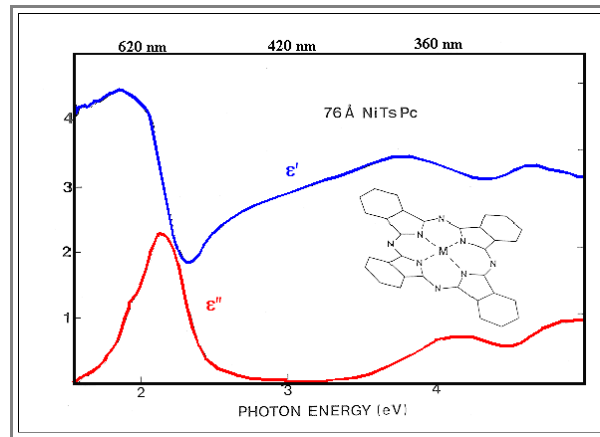


- It is just as wrong as the picture above. However, the same "excuse" applies: This is Undergraduate stuff, much simplified, and without a discussion of the statistical nature of the orientation.

Dielectric Function of a Metal-Organic Molecule

Illustration

Here is a curve of a dielectric function taken from a random scientific paper.



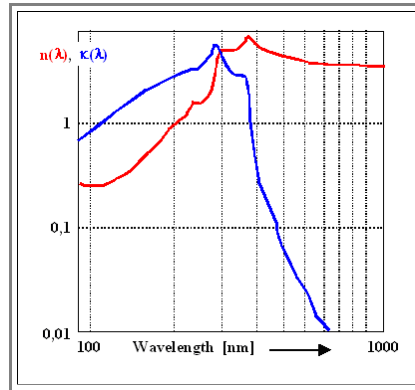
- The material is some kind of metal-organic molecule; shown in the inset.
- The real part (ϵ') and the imaginary part (ϵ'') as measured in a small range of frequencies (expressed as photon energy) is shown.

While the general structure clearly shows the effect of (a strongly damped) [oscillation mechanism](#), there is some fine structure visible. This is easily understood in principle, because all kinds of oscillations are possible in contrast to the simple model with only one oscillating atom.

Complex Index of Refraction of Silicon

Here is the complex index of refraction of *Silicon* (from Handbook of Optical Constants of Solids (E.D. Palik))

Illustration



- Note that the imaginary part κ increases as soon as the wavelength is small enough for band-to-band absorption processes; i.e. $\hbar \cdot \nu > E_G$ applied (with E_G = band gap).

Ferromagnetic Materials

Advanced

Here is a list of ferromagnetic (including **ferrimagnetic**) materials (from the "Kittel"). Besides the chemical formula, the [Curie temperature](#) T_C and the [magnetic moment](#) m_m of the molecules (in units of the [Bohr magneton](#) m_B) is listed

Material	T_C [K]	m_m
Fe	1043	2.22
Co	1388	1.72
Ni	627	0.606
Gd	292	7.63
Dy	88	10.2
CrO ₂	386	2.03
MnAs	318	3.4
MnBi	630	3.52 + 3.6 Mn −0.15 Bi
EuO	69	6.8
NiO / Fe (Ferrit)	858	2.4
Y ₃ Fe ₅ O ₁₂	560	5.0

The list could be much longer, but we see a number of interesting facts.

- There exist more **elemental** ferromagnets than just the common trio **Fe**, **Ni**, **Co** - but not at room temperature!
- Elements that do **not** form a ferromagnetic elemental crystal, may become from ferromagnetic crystals in combination with some other atoms. This is especially true for **Mn** and **Cr** compounds.
- There are many "strange" oxides or mixtures of oxides of non-magnetic elements (**EuO**) or magnetic elements that are ferri- or ferromagnetic. Most compounds with the composition **MOFe₂O₃** and **M** being some bivalent metal (including **Fe**) are **ferrimagnets**.
- The total magnetic moment carried by the atoms or molecules can be rather large; it is a combination of the moments of the atoms (and, if applicable, whatever free electrons contribute). It is constructed from positive and negative contributions for ferrimagnets (as shown for the **MnBi** case), It is, however not obvious, if the total magnetic moment from some compound is "ferro" (all contributions same direction or sign) or "ferri" (different signs). While most oxides are "ferro", some (e.g. **CrBr₃**, **EuO**, **EuS**) are "ferro".

Here is a list of **anti-ferromagnetic materials**; T_C is now the [Néel temperature](#)

Material	T_C [K]
Cr	308
MnO	116
MnS	160
NiO	525
FeCl ₂	24
FeO	198
CoCl ₂	25

CoO	291
NiCl ₂	50
NiO	525

- We find the by now usual suspects **Fe**, **Ni**, **Co**, **Mn**, **Cr** and their oxides, but also plenty of other compounds (mostly not listed).
- While anti-ferromagnetic ordering provides endless challenges for solid state physicists, it appears that there are no practical uses for this magnetic property at present.

Giant Magnetostriction

Advanced

Here are a few "clippings" concerning giant magnetostriction. (Emphasizes are mine)

Announcement of a **NATO** conference; June **2000**

The phenomenon of *magnetostriction* was discovered more than 150 years ago (Joule J.P. Philosophical Magazine, 1847, **30**, p.76). Since that time there has been both study of the basic science, and application in such areas as the generators of sound, magnetoacoustic transformers, actuators for opto-electronic systems, devices for non-destructive control and remote detection and ranging. The recent development of modern technologies, such as microfabrication, and materials, such as rare earth based bulk materials and magnetic thin films, has produced new opportunities for the study and application of magnetostriction. Thus, *discovery of giant magnetostriction*, enables one, in particular, to generate ultrasound and extend the usage of the non-destructive control techniques; development of *cryogenic technologies* gives new insight to the forced magnetostriction, namely to its irreversible component, related to *magnetization reversal* and *thermoactivated processes*, which are involved in displacement of the *domain walls and flux lines*, i.e. stability of magnetic and superconducting devices, as well as to the giant magnetostriction in *rare-earth-magnets* (up to 10^{-2}). The new field of the interest in magnetostriction as the strain derivative of *magnetic anisotropy*, is relevant to *magnetic recording industry*, particularly as recorded densities go beyond 20 Gbits/in². As physical dimensions of devices are reduced the surface area to volume ratio increases, and surface anisotropy (magnetostriction) effects may become significant in terms of ultimate switching speeds or noise floor. Miniaturisation within the sensor/actuator sector also may invoke such complications, and also now make magnetostrictive materials competitive with piezoelectric materials. There has been a resurgence of interest in perovskite materials, particularly for their outstanding magnetoresistive properties. The fundamental mechanisms driving the observed effects are still being elucidated, but lattice distortions (Jahn-Teller) and significant magnetostrictions appear to play a part. There is an urgent need for coherent studies in this area.

Magnetic field induced *giant magnetostriction* has recently been discovered in high-temperature superconductors. The magnetoelastic strains may limit technical applications of this important group of materials.

Spread of the novel experimental techniques like *magnetic resonances*, *neutron scattering*, *modern X-ray facilities* to magnetostriction examination, allows high resolution *structural studies* of magnetostriction and the differentiation of its surface and bulk components. It is timely to review and explore the various possibilities offered here, and attempt to co-ordinate the use of large scale facilities to maximise the scientific output.

The goals of the proposed ASI were delivery of lectures on new achievements and discussion of the listed potentials for the study and application of magnetostriction study among experts from the different branches of science and industry, presenting the leading teams of the West and Eastern Europe. It is hoped that a more co-ordinated and focussed approach at both the level of fundamental science and demonstrator applications, moved the subject on significantly. Wide dissemination of the meeting via publications will be an important outcome. The recent opening up of Eastern Europe makes such a meeting practical, as before much expertise lay beyond the reach of western scientists. This ASI was the first forum on modern trends in magnetostriction study and application. Only a meeting of this kind, supported by NATO, allowed us to gather the worldwide acknowledged specialists in the related fields, capable to promote the solution of the existing problems and identify the future prospects.

Dec. **2000** News from Ames Lab - Iowa State University

Giant Magnetostrictive Materials

Just about everything you ever wanted to know about magnetostriction is covered in a new book by Goran Engdahl entitled Handbook of Giant Magnetostrictive Materials.

The book is a fairly thorough treatment of giant magnetostrictive materials, from basic theory and physics to applications. The book contains six chapters and four appendices. The first chapter is entitled "Physics of Giant Magnetostriction." It covers a wide range of topics from the physical origins of giant magnetostriction to manufacturing processes. Stops along the way include metallurgy and microstructure, crystallography, atomic and magnetic force microscopy images, designing materials, crystalline rare earth alloys, applications, and, of course, TERFENOL-D, among other materials. This chapter includes 85 figures, 50 equations, and 13 tables, all of which serve as excellent illustrative guides to understanding the subject.

The second chapter is devoted to modeling giant magnetostrictive materials. Linear models, finite element modeling, and nonlinear modeling are discussed in detail. Explanations and illustrations of stress and strain, coupling, equivalent circuits, resonance, wave propagation, eddy currents, hysteresis, among others, are included. As is expected, a large number of equations are necessary to explain the topics adequately, and over 230 equations are used throughout the chapter, accompanied by 38 figures.

Chapter 3 covers magnetostrictive design and contains discussions of magnetic, electrical, mechanical, electromechanical, and thermal design. Also included are physical data of TERFENOL-D, magnetic and mechanical operation ranges, 38 figures, and 54 equations.

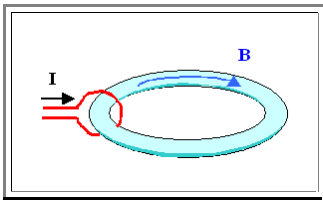
Two methods of actuator characterization using magnetostrictive materials are presented in Chapter 4. The two methods are time domain methods and frequency domain methods. The discussion is accompanied by 14 figures, 23 equations, and one table.

Device applications are covered in Chapter 5, with emphases on sound and vibration sources, vibrational control, direct and indirect motional control, and materials processing. The final chapter discusses the materials, fabrication, and application of giant magnetostrictive thin films.

Hystereses Losses

Advanced

Finding the proper formula for the hystereses losses is most easily done by considering the following situation:



- We have a single loop of wire around a doughnut of magnetic material with $R =$ (average) radius of the doughnut or torus of the magnetic material. A current I flows through the wire loop.
- The magnetic field H generated by this arrangement is given by

$$H = \frac{I}{2\pi R}$$

This formula follows straight from the [Maxwell equations](#); it is known as **Ampère's law**.

- The magnetic field H of the wire coil induces a magnetic flux B in the torus.
- If we now imagine that I changes suddenly, e.g. by ΔI in the time interval Δt , to a new constant value, the magnetic flux changes by ΔB , and a voltage U will be induced in the wire coil given by

$$U = \frac{A \cdot \Delta B}{\Delta t}$$

- With $A =$ cross-sectional area of the torus

This is of course nothing but the well-known effect of self-inductance - you cannot turn on a current very quickly that is flowing through a large inductance.

- In our "experiment", however, we just keep the current at the new constant value - even against the effect of the induced voltage that opposes current flow in the wire.
- This requires that we cancel the effect of the induced voltage by raising the outside voltage accordingly. Since we are interested in power losses, we may also argue that we now need to supply power to the system for a while to be able to keep I fixed. Note that in this kind of "experiment" we can make the wire with zero resistance, so no power is fed into the system as long as I does not change

We need to maintain a current I against a voltage U ; this requires the power $P_{\Delta B} = U \cdot I$.

- Using our formulas from above (using $I = 2\pi R \cdot H$) yields

$$P_{\Delta B} = 2\pi R \cdot H \cdot A \cdot \frac{\Delta B}{\Delta t}$$

Power is energy times time; for finding a useful material properties it is advantageous to calculate the energy E deposited in the magnetic material per unit volume

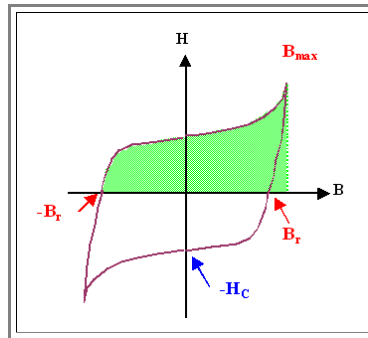
- Dividing by the volume $V = 2\pi R \cdot A$ and forming $E_{\Delta B} = P_{\Delta B} \cdot \Delta t$ gives

$$E_{\Delta B} = H \cdot \Delta B$$

The total energy deposited in an unit volume of the magnetic material during the time it takes to run through **one** cycle of the hystereses curve is obtained by integrating over a complete cycle, i.e.

$$E_{\text{cycle}} = \int_{\text{?}}^{\text{?}} H \cdot dB$$

- With slightly unclear boundaries at present. Lets look at this
- Since we are integrating over B , we rotate the hystereses curve to obtain the conventional " $y - x$ " form:



- This is exactly the hystereses curve [used before](#), we [just replaced](#) M by B which does not change the shape. For integrating once around the loop, we may start at the extreme $-B_r$ and integrate to the other end, i.e. to B_{max} .
- This gives us the area shown in green which corresponds to the energy needed to change B from B_r to B_{max} .
- Now we continue the integration running backwards from $-B_{\text{max}}$ to B_r . This gives us the small area corresponding to the green part outside the hystereses loop, however with a negative sign because we actually recover some energy stored in the magnetization of the material.

➤ In total we obtain just half of the area of the hystereses loop.

- The complete integral thus is simply the area contained in the hystereses loop.

➤ For hard magnetic materials with a roughly rectangular hystereses loop, the area and thus the energy dispersed in *one* cycle per unit volume is than approximately

$$E_{\text{cycle}} = \int H \cdot dB \approx 2 \cdot H_C \cdot B_r$$

- The total power loss than is the energy loss per cycle times the number of cycles, i.e.

$$P \approx 2 \cdot f \cdot H_C \cdot B_r$$

- With f = frequency.

➤ This was the [formula used](#) in the main part.

Comparison of Dielectric and Magnetic Properties

Here is a quick and simple comparison of dielectric and magnetic *definitions* and *laws*

Basics

Dielectric Behavior		Magnetic Behavior	
Charge q			<i>No equivalent</i>
Electrical field \underline{E}			Magnetic field \underline{H}
Electrical displacement \underline{D}			(Magnetic) Induction \underline{B}
Permittivity constant of vacuum ϵ_0			Permeability constant of vacuum μ_0
Relative dielectric constant of material ϵ_r			Relative permeability constant of material μ_r
<i>From Maxwell equations</i>			<i>From Maxwell equations</i>
Connection between dielectric flux density \underline{D} , electrical field \underline{E} , and relative dielectric constant ϵ_r	$\underline{D} = \epsilon_0 \cdot \epsilon_r \cdot \underline{E}$	$\underline{B} = \mu_0 \cdot \mu_r \cdot \underline{H}$	Connection between <i>magnetic flux density</i> \underline{B} , <i>magnetic field</i> \underline{H} , and <i>relative (magnetic) permeability</i> μ_r
Formulation with electrical Polarization \underline{P} in the material caused by the electrical field	$\underline{D} = \epsilon_0 \cdot \underline{E} + \underline{P}$	$\underline{B} = \mu_0 \cdot \underline{H} + \underline{J}$	Formulation with <i>magnetic polarization</i> \underline{J} in the material caused by the magnetic field
<i>Justified by theory of polarization mechanisms</i>		<i>Justified by theory of magnetization mechanisms</i>	
Material "law" describing \underline{P} as response of a material to a field \underline{E} and defining the dielectric susceptibility χ Note exception: <i>Ferroelectricity</i>	$\underline{P} = \epsilon_0 \cdot \chi \cdot \underline{E}$	$\underline{J} = \mu_0 \cdot \chi_{\text{mag}} \cdot \underline{H}$	Material "law" describing \underline{J} as response of a material to a field \underline{H} and defining the <i>magnetic susceptibility</i> χ_{mag} Note exception: <i>Ferromagnetism</i>
Relation between χ and ϵ_r	$\chi = \epsilon_r - 1$	$\chi_{\text{mag}} = \mu_r - 1$	Relation between χ_{mag} and μ_r
Definition of \underline{P} as material property in terms of electrical dipole moment $\underline{\mu}$ and density $\underline{N_V}$	$\underline{P} = \langle \underline{\mu} \rangle \cdot \underline{N_V}$	$\underline{J} = \langle \underline{m} \rangle \cdot \underline{N_V}$	Definition of \underline{J} as material property in terms of <i>magnetic moments</i> \underline{m} and density $\underline{N_V}$
		$\underline{M} = \underline{J} / \mu_0$	Definition of <i>magnetization</i> \underline{M}
		$\underline{M} = \chi_{\text{mag}} \cdot \underline{H}$	Relations between \underline{M} and \underline{H}

Next, let's compare mechanisms that lead to polarization

Dielectric Polarization		Magnetic Polarization	
<i>Electronic polarization</i>			<i>Diamagnetism</i>
Induce dipole moments by displacing electrons and nuclei. Weak for spherical atoms. Stronger for covalent bonds. Important for optics.	$\epsilon_r \approx 1,0001 \dots 30$	$\mu_r \approx 0,9999$	Induce precession of electrons. Always very weak and opposite to field. Not important.
<i>Orientation polarization</i>			<i>Paramagnetism</i>
Average small orientation of fluctuating existing dipoles. Only in <i>liquids</i> ; can be large. Not important.	$\epsilon_r \approx 2 \dots 100$	$\mu_r \approx 1,0001$	Average small orientation of existing dipoles free to rotate in <i>solids</i> . Always small; not important. Extreme case: <i>Ferromagnetism</i> .
<i>Ionic polarization</i>			No direct counterpart

Net dipole moment from distribution of charges. Important.	$\epsilon_r \approx 2 \dots 100$		
<i>Ferroelectricity</i> Natural dipoles defined by crystallography are lined up. Important.	$\epsilon_r > 1000$	$\mu_r > 1000$	<i>Ferromagnetism</i> Natural magnetic moments are lined up in any directions (with crystal directions preferred). <i>Extremely</i> important.

Multiple Choice Test zu

4.1.2 Origin of Magnetic Dipoles

Start Multiple Choice

Multiple Choice Test zu

4.1.3 Classification of Magnetic Materials

Start Multiple Choice

Multiple Choice Test zu

4.1.4 Definitions and General Relations - Summary

Start Multiple Choice

Multiple Choice Test zu

4.3.3 Magnetic Domains

Start Multiple Choice

Multiple Choice Test zu

4.3.4 Domain Movement

Start Multiple Choice

Multiple Choice Test zu

4.3.5 Losses and Frequency Behaviour

Start Multiple Choice

Multiple Choice Test zu

4.3.7 Ferromagnetism - Summary

Start Multiple Choice

Multiple Choice Test zu

4.4.1 Application of Magnetic Materials

Start Multiple Choice

Multiple Choice Test zu

4.5.1 Magnetic Materials - Summary

Start Multiple Choice

Exercise 4.3.1 Maximal Magnetization



What is the ultimate magnetization we can reasonably expect for a permanent magnet?

- Consider reasonable magnetic moments of atoms in combination with reasonable densities
- Compare to the graph given in the backbone.


Illustration



Link to the [solution](#)

Exercise 4.3-2: Magnetic Moments of Fe, Ni, Co

 Given the type of the lattice, the lattice constants of **Fe**, **Ni**, **Co** ([look it up!](#)), and the magnetization curves in [chapter 4.3-2](#), the question is:


-  How large are the magnetic moments of these atoms in terms of a Bohr magneton?

Illustration



Link to the [solution](#)

Solution to Exercise 4.3-1

 Surprise! No solution given here!
Do it yourself! It is quite illuminating.

Illustration

Solution to Exercise 4.3-2

Illustration

Given the type of lattice, the lattice constants of **Fe**, **Ni**, **Co** ([look it up!](#)), and the magnetization curves in [chapter 4.3-2](#): How large are the magnetic moments of these atoms in terms of a Bohr magneton?

Simple - but still a bit tricky.

First we get the basic data:

- Lattice **Fe**: **bcc**; lattice constant **a** = 2.86 Å; atomic density $\rho_A(\text{Fe}) = 2/0.286^3 \text{ atoms/nm}^3 = 85.5 \text{ atoms/nm}^3$
- Lattice **Ni**: **fcc**; lattice constant **a** = 3.52 Å; atomic density $\rho_A(\text{Ni}) = 4/0.352^3 \text{ atoms/nm}^3 = 91.7 \text{ atoms/nm}^3$
- Lattice **Co**: **bcc**; lattice constant **a** = 2.51 Å, **c** = 4.07 Å; atomic density $\rho_A(\text{Co}) = 2/[1/2 \cdot c \cdot a^2 \cdot 3^{1/2}] \text{ atoms/nm}^3 = 90.1 \text{ atoms/nm}^3$

Then we realize that the curves in [chapter 4.3-2](#) give the maximum magnetization, i.e. the magnetization state for all magnetic moments perfectly aligned. From the figure we can deduce the following numerical values for the saturation magnetization m_{Sat} :

- $m_{\text{Sat}}(\text{Fe}) = 17 \cdot 10^5 \text{ A/m}$
- $m_{\text{Sat}}(\text{Ni}) = 5 \cdot 10^5 \text{ A/m}$
- $m_{\text{Sat}}(\text{Co}) = 14 \cdot 10^5 \text{ A/m}$

However, the units shown are **A/m**, which are not what we would expect. Obviously we must convert this to - well, what exactly?

If we look at a Bohr magneton, m_{Bohr} , we have

$$m_{\text{Bohr}} = 9.27 \cdot 10^{-24} \text{ Am}^2$$

Obviously, the unit we need is **Am²**. We obtain that by multiplying the **A/m** by **m³**, which makes clear that the m_{Sat} numbers given are per **m³** - as they should be!

The magnetic moments m_A per atom are thus

$$m_A = \frac{m_{\text{Sat}}}{\rho_A}$$

What we obtain is

$$\begin{aligned} m_A(\text{Fe}) &= \frac{17 \cdot 10^5 \text{ A/m}}{85.5 \text{ atoms/nm}^3} = \frac{17 \cdot 10^5 \text{ A} \cdot 10^{-27} \text{ m}^3}{85.5 \text{ m}} = 1.98 \cdot 10^{-23} \text{ A/m}^2 = 2.14 m_B \\ m_A(\text{Ni}) &= 5.45 \cdot 10^{-24} \text{ A/m}^2 = 0.588 m_B \\ m_A(\text{Co}) &= 1.55 \cdot 10^{-23} \text{ A/m}^2 = 1.67 m_B \end{aligned}$$

Now that is an interesting result! It's satisfying because we actually get sensible numbers close to a Bohr magneton, and it's challenging because those numbers are not very close to 1, 2, or possibly 3.

For example, how can a **Ni** atom have a magnetic moment of **0.588 m_B**, and a **Fe** atom one of **2.14 m_B**, considering that the spins of the electrons carry exactly **1 m_B**?

There are two possibilities for this apparent discrepancy:

- Our calculation is somehow a bit wrong
- There are some effects not yet discussed that change the magnetic moment an atom in a crystal lattice carries around with itself somewhat.

The first possibility can be ruled out, because in standard textbooks, e.g. in the "*Kittel*" we find the following values for m_A

- $m_A(\text{Fe}) = 2.22 m_B$
-

- $m_A(\text{Ni}) = 0.606 \text{ m}_B$
- $m_A(\text{Co}) = 1.72 \text{ m}_B$

Not identical, but close enough. In fact, looking more closely, the Kittel values are for $T = 0 \text{ K}$, whereas our values are for room temperature $T = 300 \text{ K}$ and thus should be a bit smaller.

- Obviously, this leaves us with some effects not yet discussed. What these effects could be, we can only guess at. Here is a short list:
 - There might be some interaction between the spins of the electrons and the "orbits" of the electrons that modifies the magnetic moment
 - The free electrons of the electron gas in our metal also "feel" the ordered spins of the atoms and react to some extent by adjusting their spins.
- This can lead to quite sizable effects. Dysprosium (**Dy**), for example, a rare earth metal, is a ferromagnet below its Curie temperature of **88 K** and its atoms then carry an $m_A(\text{Dy}) = 10.2 m_B$.

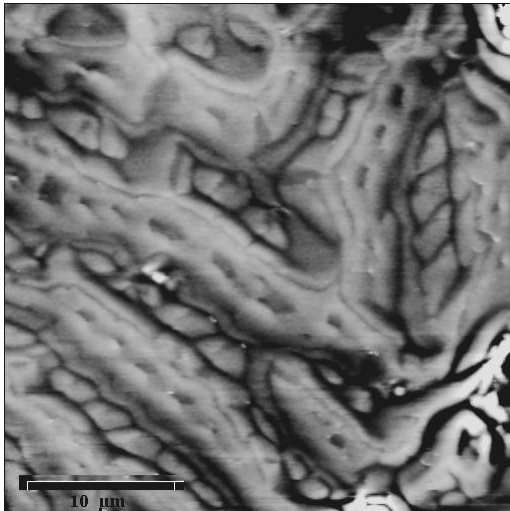
Domain Structures

Here are some domain structures as found in the Web. These pictures are scientific citations and thus do not infringe on Copyrights.

Unfortunately, some of these pictures do not obey the first law of scientific picture publishing, which is:

You must always have a scale on your micrographs!

Illustration



Domain structure of native *magnetite* (obtained by "magnetic force microscope imaging"). It is quite complicated.
Source: <http://premare.imr.tohoku.ac.jp/SDurbinWWW/magndomn.html>.

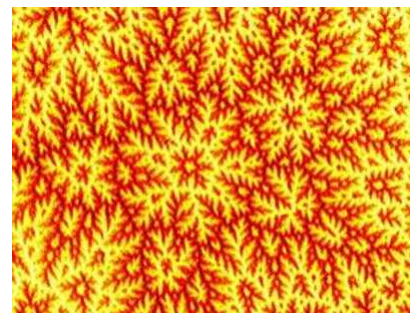
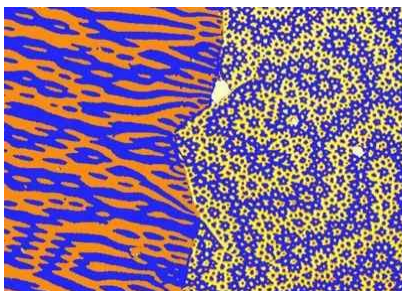


A very nice picture of the domain structure of **Fe + 3% Si**.

(From a paper entitled: Interaction Between domain walls and structural defects as a background for magnetic inspection of steel structures" by E.S. **Gorkunov**, Yu.N. **Dragoshansky** (from the Proceedings of the 15th World Conference on Nondestructive Testing Roma (Italy) 15-21 October 2000))




A labyrinth of magnetic domain as seen in an amorphous membrane of **Gd-Co**.
Historical picture from unclear source.



Domains in **NdFeB** material. The domain character in this highly anisotropic material depends on the orientation of the easy axis of magnetization relative to the surface. In the grain on the right the axis is oriented perpendicularly, resulting in fine branched domains, while the grain on the left has a lower degree of misorientation relative to the surface.

From A. Hubert and R. Schäfer, *Magnetic Domains*.

The Analysis of Magnetic Microstructures, Springer, Berlin-Heidelberg-New York, 1998. With friendly permission of R. Schäfer.



Domain branching observed at the surface of a **Co** crystal with strong out-of-plane anisotropy.
Same source as to the left.

Permanent Magnets

Below is one table with some data:

Illustration

Permanentmagnete			Relat. Energieprodukt	Relat. Koerzitivfeldstärke	Arbeits-temperatur	Curie-temperatur	Anmerkungen
	Ferite	Strontiumferrit	1	~3,1	< 250 °C	~450 °C	Wenig kräftig, hart und spröde, leicht entmagnetisierbar
		Bariumferrit	~1,1	~4,8			
	Aluminium-Nickel-Cobalt	AlNiCo (500)	~1,4	1	< 400 °C	~890 °C	Wenig kräftig, sehr leicht entmagnetisierbar
Seltene-Erde-Magnete	Samarium-Cobalt	SmCo ₅ (18)	~6	~13,5	< 250 °C	~720 °C	Kräftig, hart und spröde
		Sm ₂ Co ₁₇ (24)	~8	~14,8		~800 °C	
	Neodym-Eisen-Bor	NdFeB (35)	~11	~17	< 120 °C	~310 °C	Sehr kräftig, hart und spröde, oxidieren leicht

Das Energieprodukt ist ein Gütewert. Die Koerzitivfeldstärke gibt die Feldstärke an, die einen Magneten entmagnetisiert.

© <http://www.wundersamessammelsurium.de/Magnetisches/XtraPermanent/index.html>

- I intended to add more information but got bored with the subject and quit. That doesn't mean that permanent magnets are not quite important. They are.

The Polyimide Story

Here is a (slightly edited) article about the development of polyimides, which appeared in 2005 in the MATERIALS SCIENCE AND TECHNOLOGY NEWSLETTER.

It is one of the rare articles that also mentions what went wrong. It gives a good idea how difficult it is in the real world out there to introduce a new material in a complex product.

Here is the [link to the Journal](#);

Copyright courtesy of **Dr. Lacombe**, the author of this story.

Advanced

MATERIALS SCIENCE AND TECHNOLOGY NEWSLETTER

Vol. 2, No. 2

Dr. Robert H. Lacombe
Chairman
Materials Science and Technology
CONFERENCES, LLC
3 Hammer Drive
Hopewell Junction, NY 12533-6124
Tel. 845-897-1654
FAX 212-656-1016
E-mail: rhlacombe@compuserve.com

EDITORIAL NOTES

During the week of November 7, 2005 MST CONFERENCES will be holding the 4th in its series of symposia on Polyimides and High Temperature Polymers along with the 2nd International Symposium on Adhesion Aspects of Thin Films in Savannah Georgia. Apropos of this event, this issue of the Newsletter will be focusing on these two topics and their interrelation. The following narration gives some reminiscences of our experiences with these topics as they played out in the mainframe computer business during the late 1970's and 1980's.

Polyimides, Stress and Adhesion: or There are no Bad Materials, Just Bad Uses of Good Ones:

Prelude to polyimides as electrical insulators

Subsequent to the DuPont company's patenting of the polyimide materials, one of the first major applications was for electrical insulators capable of sustaining high temperature conditions. To those of us who were first exposed to the properties of engineering thermoplastics in the 1950's, the polyimide materials were a marvelous revelation. Those who recall the early days of manufactured goods made of plastic materials will remember that if subjected to high temperatures, say the flame from a match, these items would either melt into a viscous goo or possibly also ignite and burn. In addition, articles made from the early resins were highly prone to fracture as in the case of children's toys made from polystyrene. On the positive side, however, the early plastics were very easy to process into a large variety of useful shapes by either molding or casting methods. Thus one could easily and very cheaply manufacture useful household items such as knives, forks, cups and spoons. So cheaply in fact that these became throwaway items. One could even get around the brittleness problem by using polycrystalline materials such as polyethylene or polypropylene. These resins could be drawn or extruded into flexible sheets or tubing for use in packaging or plumbing applications. However, these materials also had the problem of temperature stability and would either distort or melt at elevated temperatures at or above 100C.

Thus it seemed that the price one paid for low cost and ease of manufacture was limited thermal-mechanical performance. Now, however, along came the polyimide materials which were clearly polymeric in nature but which also possessed extraordinary thermal stability previously unheard of in a polymer. The earliest polyimide to be produced in commercial quantities was the KAPTON[®] (The basic resin goes under the name pyromellitic dianhydride oxydianiline or PMDA-ODA for short) film material produced by the DuPont company. Here was a flexible plastic film that you could take a match to and it would not burn or melt. Under a nitrogen environment this material could be shown to be stable up to 400C in a Thermo-Gravimetric (TGA) experiment. This was truly astounding behavior for a polymer and it was felt that here was a class of materials that could surmount the thermal-mechanical limits of the ordinary thermoplastics. However, high performance came at a price. It turned out that in order to process the KAPTON material one had to use concentrated sulphuric acid or some other equally potent solvent. It

seemed there would be no free lunch. If you wanted the high performance thermal-mechanical properties you had to sacrifice ease of processing.

All was not lost though since there was another way to make polyimide films without resorting to extreme processing techniques. It turns out that nearly every polyimide has a twin sister known as a polyamic acid. The polyamic acid has a chain architecture very similar to its polyimide sibling except that instead of the refractory imide group one has an acid moiety which gives the chain much more flexibility and also makes it much more soluble in relatively benign solvents such as N-methy-pyrrolidone (NMP). Thus one first casts a film of the amic acid material out of NMP and then after drying the amic acid can be converted to the polyimide by curing at an elevated temperature. With this processing strategy it seemed as though the ultimate miracle had finally been achieved. Now one had a relative simple and benign process for achieving high performance films and coatings.

It was at this stage in the late 1960's and early 1970's that the microelectronics industry started to take a serious interest in using the polyimide materials as high performance insulators for semiconductor devices and advanced multichip modules. Up to that time the insulating material of choice was sputtered glass. This material is essentially amorphous SiO_2 which is the purest form of glass known and forms the basis of nearly all glass artifacts. The glassware in your cupboard is basically SiO_2 doped with boron and other materials which are used to give the final artifact a number of desired properties such as color, density, ... etc. SiO_2 was also a natural for the microelectronics industry since nearly every silicon wafer comes with a thin layer of SiO_2 due to the natural oxidation behavior of the pure silicon. This thin oxide layer is the primary reason that silicon is the primary semiconductor material used for microchips in the world today even though one can in principle use materials such as gallium arsenide to make significantly faster devices. gallium arsenide does not come with a built in insulating layer whereas the oxide layer on the silicon serves as a convenient and highly efficient built in insulator which is crucial to the manufacture of integrated circuits. All circuits require three separate materials: a semiconductor, a conductor and an insulator each of which is critical in fabricating the final device. By chance SiO_2 came as a made to order insulator on top of every silicon wafer which could be used as is or enhanced in thickness for wiring up transistors and other devices in integrated circuits.

Amorphous SiO_2 was indeed a wonderful insulator material but it came with its own set of drawbacks which became painfully apparent as soon as one wanted to make multilevel structures. Among the list of problems with using SiO_2 were the problems that such coatings were subject to pin hole defects if they were too thin and if too thick they were susceptible to cracking also. However, one of the biggest problems was the fact that SiO_2 coatings were conformal, i.e. they closely replicated the underlying surface. This is a killer issue when making multilevel circuits. The first level is no problem since one is building on top of a flat silicon wafer. After the first layer is built it has a complex topology of metal lines and other structures which need to be covered over before the next layer can be built. This is where the problem with SiO_2 starts, since after coating, all of the hills and valleys of the underlying circuitry will be replicated leaving a lumpy surface. This surface must somehow be flattened before the next layer can be built. One would much rather have a coating that filled in all of the hills and valleys leaving a completely flat surface the same way that water filling a rough basin leaves a perfectly flat surface regardless of the topology of the underlying lake bed. In the jargon of the industry what one wants is a planarizing coating material that will leave behind a flat smooth surface after coating.

Enter the polyimide materials. Since these materials go down as a viscous liquid, they tend to planarize the underlying topology much better than SiO_2 . The degree of planarization is not perfect but is nonetheless much much better than what one gets with SiO_2 . Thus it looked as though the polyimides were a natural to replace SiO_2 as the insulator of choice for multilevel circuits, since not only did they planarize much better they also had a significantly lower dielectric constant which for the circuit designers meant that one could go to higher wiring densities before the problem of crosstalk between metal lines became a problem. Thus it was that in the early to mid 1970's very serious development efforts got underway within the microelectronics industry to use the polyimides as replacement materials for SiO_2 in multilevel wiring structures. It was also at this time that the staff of MST got their first exposure to the polyimides and the host of problems that came along with them.

Professor J. E. Gordon has neatly summed up the nature of dealing with new materials with this insightful comment: "A deep, intuitive appreciation of the inherent cussedness of materials and structures is one of the most valuable accomplishments an engineer can have. No purely intellectual quality is really a substitute for this" ("Structures or why Things Don't Fall Down", J. E. Gordon (Da Capo, Press, 1978) p. 63). We were about to learn the truth of this of this statement firsthand as we started upon a development program to use the polyimide materials as insulator layers in multilevel wiring structures. But first a little introduction to the essential architecture of mainframe computers.

Wiring needs of a mainframe computer

All digital computer devices are what might BE called "vertically integrated". That is they form a layered

structure. At the lowest level, the basic calculations and data manipulations are carried out by transistors, diodes and resistors on silicon chips at the scale of roughly 1 micrometer and for the most advanced devices less than half that dimension. This is referred to as the chip level. A problem immediately arises due to the fact that humans cannot directly read data stored at the 1 micrometer scale. The digital signals must be scaled up to a coarser size or in technical jargon fanned out in order to be interpreted. The way this is done is to package the silicon chip on a larger substrate, typically but not necessarily a ceramic block, where the wiring structure is fanned out to roughly 100 times the dimension that exists on the chip. This is the second layer commonly referred to as the second level of wiring or the first level of packaging. If this is still not enough, yet a third level of wiring is employed in a third layer whereby the ceramic block is plugged into an epoxy based board which fans out the wiring density another 10 to 100 times. By this stage one usually achieves a wiring density that can be dealt with by human fingers which can then plug the wires into useful devices such as disk drives, display screens and keyboards. The polyimide materials were being considered for application at both the first and second level of wiring and this is where the story starts to get interesting. First, however, we need to digress briefly on the overall architecture of mainframe computers.

In the late 1960's and early 1970's the IBM company was having great success packaging single chips on ceramic substrates roughly 3/4" square. These ceramic modules would then be plugged into epoxy boards which could further be mounted into racks which would then constitute a mainframe computer. This was all well and good but there was still much room for improvement. In the first place all those cards and boards mounted in racks were large messy affairs taking up a lot of room and requiring extensive support structures such as cooling fans and power supplies. Another even more serious problem was the fact that by mounting each chip on a single ceramic module one was setting up a communications barrier of sorts between the chips. For example, say that a central processing circuit on one chip needs a piece of data on a remote memory chip. The electrical signal which transfers the data has to go from the chip to the ceramic module, thence to the supporting card, from there to the ceramic module supporting the memory chip and then to actual chip itself which sends the desired data back by the same route. It quickly became apparent to the electrical designers that the main bottleneck standing in the way of improving machine speed was the highly spread out packaging structure which single chip modules required. One idea that came up was to make the machine as one huge integrated circuit on a 3 inch wafer. This never flew for a number of reasons not the least of which was the problem of doing 1 micrometer lithography accurately over a span of 3 inches. The next level of thinking said that if we cannot cram the whole machine onto a single wafer at least we can mount all the chips on a single substrate and thereby eliminate one full level of packaging. At this stage the concept of a multichip module was born. Not only that, this concept was vigorously implemented resulting in multichip ceramic modules with approximately 40 levels of wiring buried in the ceramic and supporting up to 100 chips. It is hard to convey just how successful this multichip module concept was. By eliminating one full level of packaging the IBM company was able to manufacture machines in the early 1980's using chip technology that had been fully developed and amortized by the mid 1970's. These machines were cheaper than and out performed those of competitors which were using the latest expensive chip technology in order to achieve similar performance. I think one can safely say that this technology made a major contribution to the roughly \$4billion/year profit that the company was earning in those days.

Implementing polyimides: or development programs always work out better in slide presentations than in the development lab

Part one: prelude to selecting a material

If it is true that in the fashion industry one can never be thin enough and in business never rich enough then in the computer industry ones machines can never run fast enough. Already in the late 1970's we were looking for ways to improve the performance of the large multichip substrates. The material being used at that time was essentially alumina (Al_2O_3) ceramic. This was the same material used for the single chip modules and was an industry standard insulator for all applications requiring stability at high temperatures. Its main strengths are its excellent insulating properties and outstanding thermal-mechanical durability. Its main weakness is its relatively high dielectric constant and its highly refractory nature which forces the use of equally refractory metals such as molybdenum as the embedded electrical conductor. This arises from the fact that the alumina and the chosen metal must be co-fired together in a furnace in order to achieve the final multilayered structure. Since firing alumina requires temperatures on the order of 2000 C only the refractory metals such as molybdenum can be used. Other metals such as copper and aluminum simply cannot take the heat!

From the electrical design point of view this was a double whammy both in terms of achievable signal speed and wiring density. As a rule of thumb the fastest and densest circuits are achieved by using a low dielectric constant insulator combined with a high conductivity metal. The alumina/molybdenum pair were quite mediocre performers on both counts. However, a solution to this dilemma quickly presented itself due to the dual electrical functions performed by the packaging substrate. At the most basic level the substrate has to provide power for all of the chips. The power signal does not have to be exceptionally

fast and the requirement is readily met by the alumina/molybdenum combination. At the next level comes the signal wiring which must transfer data among all of the chips. The machine performance is critically dependent on the speed of this wiring and this is where alumina/molybdenum fails to make the grade. The resolution to the problem now pops out. Make the basic substrate out of alumina/molybdenum to handle the power requirements and then build thin film polyimide/metal wiring on top to handle the signal wiring where the metal can now be a high conductivity material such as copper or aluminum.

The stage was now set for the polyimide materials to make a major contribution to high level packaging technology. All that had to be done was to select the appropriate material and make it work. In all of the group meetings and slide presentations it looked like a done deal. However, the cussedness of the real material world was about to enter the stage in a most decisive way.

The first step of course was to choose an appropriate polyimide material. This is a far more critical step than might be appreciated at first. There are essentially a near infinite number of polyimides to choose from depending on the details of the backbone architecture and any side groups that might be added. Nonetheless, a choice had to be made and made quickly since in the microelectronics industry any development program has to "hit the ground running" as the paramilitary people would put it. Requirement number one was that whatever material was going to be used had to be commercially available, which greatly pared down the field of choices. The final choice was made even easier since there was one polyimide that had already been experimented with at the chip level in making 1 micrometer thick insulator layers. This was a commercial product that went under the name "Skybond 703"[®]. The original application was apparently as a high temperature glue for high performance fighter aircraft. This material had all the desirable properties required of an insulator material including a low dielectric constant but most importantly it planarized very well. What cinched the case for the packaging laboratory, however, was the fact that this material already had a track record of sorts. Thus the program went forward to fabricate thin film wiring onto of ceramic substrates using the Skybond material. Everything looked fine since the only change being made was that 10 to 15 micrometer coatings would have to be fabricated instead of 1 micrometer layers which the chip people were using. Now how could simply making thicker coatings cause any problems? (Answer: The tendency of coatings to crack and delaminate scales linearly with the coating thickness. This comes directly out of elementary fracture mechanics)

Part two: revenge of the material gods

In retrospect it is hard to believe that one could have found a worse material for the application at hand. In essence had we systematically screened all of the polyimide materials and then on the basis of a rigorous analysis tried to select the worst possible material I doubt that we could have beaten the Skybond material. As far as I can remember not a single viable substrate was made. The coatings all had massive cracking problems due to a number of thermal-mechanical problems that went entirely overlooked. The first problem was that Skybond had nearly all the mechanical properties of window glass. The strain at break was less than 2% and if you looked at a typical stress-strain curve and suppressed the scales on the axes the diagram was indistinguishable from that which a glass sample would give. This by itself does not necessarily imply disaster since many glue type materials are mechanically brittle but still perform adequately. However, the second problem that descended was the fact that the thermal expansion of the Skybond material was some 10 times larger than that of the ceramic onto which it was coated. This should have set off alarm bells since, as mentioned above, the polyimide materials have to be cured at an elevated temperature which in the case of Skybond was roughly 400 C. Oh Oh, could be trouble ahead. A little arithmetic immediately reveals that on cooling from the curing temperature at 400 C to room temperature at 20 C the part in question will have undergone a thermal excursion of roughly 380 C which, when multiplied by the thermal expansion difference between the polyimide and the ceramic, gives rise to a rather substantial thermal strain. There is a handy little equation called the membrane formula which allows one to quickly estimate the expected stress level in a coating which is subjected to these process conditions. In a nutshell it goes like this:

$$\sigma = E \times a \times \Delta T / (1 - \nu)$$

In this equation E is the modulus of the polyimide, thermal expansion difference between the polyimide and the ceramic, T the temperature excursion and Poisson's ratio of the polyimide. In this example $E = 3 \text{ GPa}$, $\nu = 1/3$, $a = 30 \times 10^{-6} / \text{C}$ and $\Delta T = 380 \text{ C}$. Inserting these values into our little formula gives an expected biaxial stress of $\sigma = 0.051 \text{ GPa}$ (GigaPascal) or in more convenient notation 51 MPa (Megapascal). Now what does this value 51 MPa mean? Looking at a stress strain diagram for Skybond one immediately notices that the ultimate tensile strength of the material is close to 70 MPa. Thus upon coating Skybond onto a ceramic substrate and curing at 400 C one has a uniform coating that is already stressed to about 3/4 of its ultimate breaking strength. Not enough to cause a problem you say. This is true. The properly fabricated blanket coatings almost never gave a problem. Unfortunately one cannot stop at a blanket coating. One also has to drill via holes in the film and fill them with metal in order to electrically connect to the wiring below. However, by doing so one creates what is known as a stress riser or point of stress concentration. A fairly elementary analysis demonstrates that close to such a hole the local stress field

increases by a factor of 2 to 3 which automatically puts the stress level over the top for Skybond. (A detailed discussion of the stress state near a via hole can be found in the following: "Stresses in thin Polymeric Films: Relevance to Adhesion and Fracture", Robert H. Lacombe in Surface and Colloid Science in Computer Technology, Ed. K. L. Mittal (Plenum Press, New York, 1987) p.179) This was in fact what was observed in all cases. As soon as the parts emerged from the plasma etcher where the via holes were bored, ubiquitous radial cracks were observed emanating from nearly every via structure. In fact nearly any kind of imperfection in the coating such as particle contamination or a defect in the substrate would also act as a stress riser and thereby give rise to cracking of the polyimide.

Part 3: revenge of the adhesion gods

OK, so we learned that it was not a good idea to go with a brittle material and maybe also we would have to pay more attention to the thermal mechanical properties of the polyimide as well as its planarizing properties and whether or not it is commercially available. Quite after the fact I looked into the history of the development process and thereby came across a memo describing the material property selection criteria. In this document was a table listing about a dozen different polyimides along with columns specifying their electrical properties, how well they planarized, whether they were commercially available ... and so on. There was only one column listing a thermal-mechanical property and that was the thermal expansion coefficient. However, the memo went on to point out that, since all of the polyimides listed had roughly similar thermal expansion coefficients, this property would thereby not be considered. Oh well, at least someone was starting to worry about the thermal-mechanical properties of the polyimide but somehow could not quite make a case for paying close attention to this detail. This was hardly surprising since in the early days of the program the entire team consisted of either chemists or electrical engineers who were really not in a position to appreciate the subtleties of stresses in thin films.

It quickly became clear that using a brittle material was not going to work and as soon as someone familiar with the mechanical properties of polyimides was brought on board it was pointed out that there are mechanically tough materials available. In fact the PMDA-ODA material which forms the basis of the KAPTON films mentioned above was one such material. Unlike the other polyimides which were essentially amorphous materials the PMDA-ODA material was quasi crystalline in that it had an ordered liquid crystalline type of morphology due to the length and rigidity of the backbone repeat unit. The stress strain diagram of this material looks almost like a rubber with a strain at break approaching 100% or better. It is the semi-ordered morphology that explains the toughness of the material. If you have an amorphous glassy material like the standard polyimides, and you apply a tensile load to it, there is not much that the material can do to accommodate the applied load. The induced stress in the material must be supported by the inter-chain van der Waals attractive forces and these are not very strong and are very short range to boot. Thus one gets a very low strain at break. With the PMDA-ODA material, however, the initial effect of the applied load is to start pulling apart and stretching out the semi-crystalline phase and only after this is accomplished does the load shift entirely to the inter-chain bonding forces. This effect is most dramatic in polycrystalline polymers such as polyethylene which can stretch out to 200% or better.

Thus, by going over to the PMDA-ODA material the problem of mechanical cracking was eliminated but unfortunately the thin film stresses remained since PMDA-ODA has roughly the same thermal expansion behavior as the other polyimides. The immediate consequence of this was that instead of cracking the material delaminated. The gods of adhesion now extracted their revenge. Coating PMDA-ODA on nearly anything such as silicon or SiO₂ tends to give poor adhesion due mainly to the chain rigidity, limited scope for chemical interaction and the pervasive presence of high residual stress levels. Thankfully, the chemists were now able to step in and alleviate this problem with the use of coupling agents such as the silane materials. The silanes improved the adhesion of PMDA-ODA tremendously and allowed the program to progress to the point where thin films with full via structures could be fabricated.

Part 4: Twilight of the polyimide program

I would like to report that at this stage the program went forward ending in a triumphal success. However, this was not to be. Development programs involving advanced materials technologies are complex and chaotic processes and the final result tends to be determined as much by luck and circumstances as by engineering skill and design expertise. Too much time and resources had been dissipated with too little in the way of results coming out of it. In addition, an even more ominous threat came onto the scene. The ceramics group which had developed and implemented the alumina based technology was not standing idly by. They essentially came upon a ceramic material which could be sintered at under 1000 C and thus could be co-fired with copper. Not only that, this material also had a low dielectric constant. Thus coupling a low dielectric constant with a high conductivity metal the electrical designers could easily show that an all ceramic/metal module was possible with greatly improved signal speed and wiring density. Two other circumstances, however, were what really sealed the fate of the polyimide program. The major circumstance was essentially political. The ceramics group

was essentially the party in power having already developed and implemented the alumina/molybdenum technology and as a consequence held most of the positions of authority and clout within the organization. The second circumstance which provide the "coup de grâce" was the fact that the limited resources available allowed only one program to go forward. Thus ended the polyimide program.

Epilog

As fate would have it the material and adhesion gods were just beginning to show their wrath and were now about to truly demonstrate the cussedness of real materials. Whereas it was true that the new ceramic material had much better electrical properties than alumina, its thermal-mechanical properties were a disaster, even worse than Skybond. In order to achieve a low sintering temperature the new material had to have a high glass content and this essentially gave it all the properties of window glass. Worse yet, as is well known to all those who work with glass, it had very poor adhesion to copper. As a consequence cracking and delamination problems descended with a vengeance. In retrospect all of this could have been anticipated in advance but one always has to remember the chaotic nature of real world development programs. Time, resources and expertise were in very short supply and the need to press forward unrelenting. Leaving out the gruesome details, the ultimate consequence of these technology failures were truly tragic. For the multichip module program the critical upshot was that a technology which should have been delivered in 1985 did not materialize until approximately 1990. This is a 5 year delay which in the microelectronics industry amounts to an eternity. One can never afford that kind of delay. While the development lab was wrestling with the demons unleashed by the material and adhesion gods, microprocessors and personal computers based on them were steadily advancing. These devices were selling for a few thousand dollars whereas mainframe computers were going for closer to a million. In order to be competitive the mainframe had to offer much greater computing power. A factor of 2 or 3 was not good enough since PCs were so cheap one could just let a few of them run overnight. No, the mainframe had to offer a bare minimum of at least a factor of 10 to 100 greater compute power. In essence the only way a customer is going to shell out mainframe level dollars is if the mainframe would handle work loads the PC's could not. One had to deliver what is called an "enabling" technology. One that could deliver what the competition could not. If we could have delivered the advanced multichip module technology in 1985 as planned then mainframes 100 times faster than PC's would have emerged. Delaying the technology for 5 years made that dream impossible.

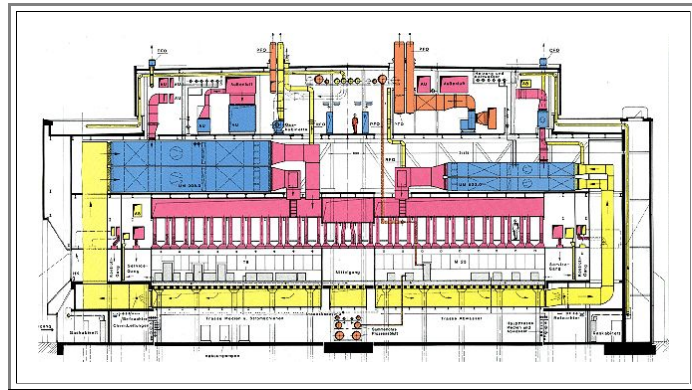
Ending on a more upbeat note, the polyimide materials have nonetheless done rather well in the electronics industry as a whole. In particular they have found widespread use in flex circuitry whereby integrated circuit modules are mounted directly onto a flexible polyimide film similar to the KAPTON material mentioned previously. All of the wiring is electrodeposited copper which can be made to adhere very well to polyimide and, since everything is flexible, problems with cracking are avoided as well. This technology has found widespread use in everything from CD Players to ink jet print heads.

In closing it is appropriate to reiterate what is I believe the universal theme of this essay "there are no bad materials only bad uses of good ones". All materials are essentially indifferent to our needs and aspirations, they obey the laws of physics and chemistry very rigidly and that is that. Any material whatever can find successful application so long as we understand its inherent nature and physical limitations. The Skybond material works quite well as a high temperature glue so long as you understand its limits. Developing such an understanding for polyimides and other high temperature polymers is one of the major goals of the upcoming MST symposia on HIGH TEMPERATURE POLYMERS and ADHESION ASPECTS OF THIN FILMS.

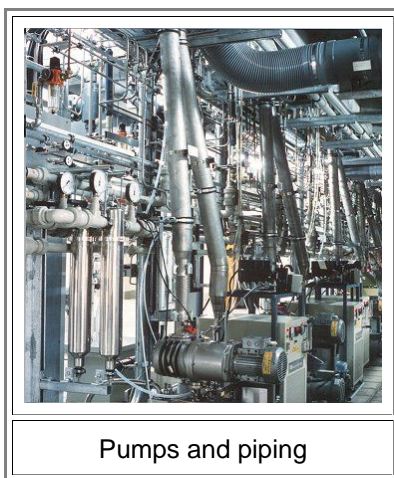
Cleanrooms

Advanced

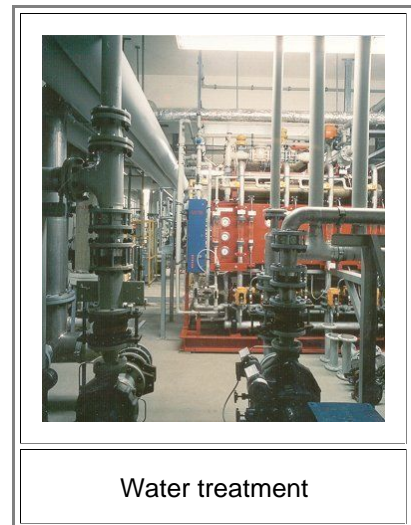
- First, let's look at the cross-section of a typical "class 1" cleanroom. Class 1 means roughly that there will be at most 1 particle per **foot³** (about 30 liters) larger than 0,2 μm or so in the air.



- If you click on the image, you will get a real big (702 kb) version where all the details are shown. Anyway, even in the small illustration you can see that the "actual" cleanroom where people make chips, is a small part of the building (the whitish portion just above the lower yellow part).
 - Everything colored is just for moving air around, keeping its temperature and humidity constant, add some fresh air from the outside and to get rid of "spent" air.
- A particular interesting place in a cleanroom building is the "basement" right under the actual cleanroom. It houses a large part of the "equipment", e.g. pumps, liquid and gas inlets, outlets, and cleaning parts, transformers, power equipment, heaters etc. It also houses miles of tubing for delivering away and taking gases and liquids. Some pictures:

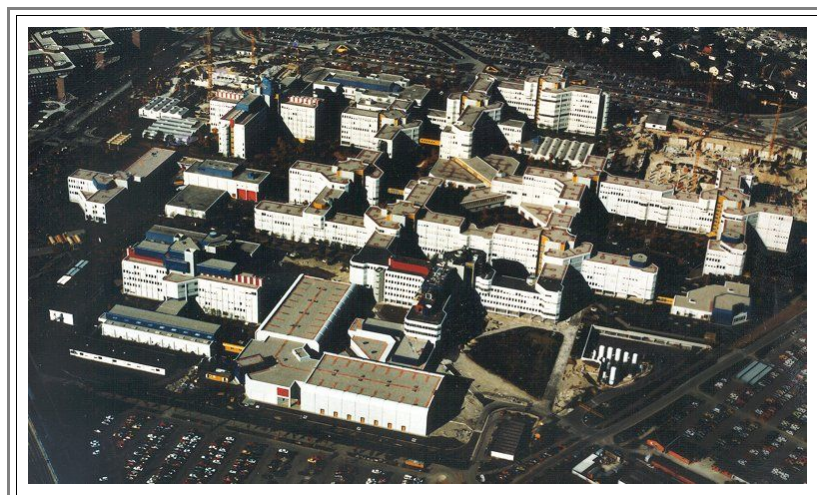


Pumps and piping

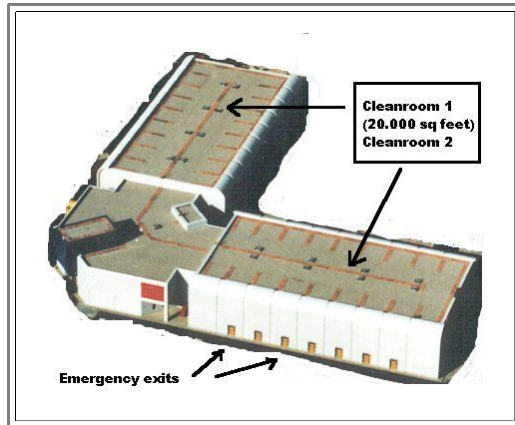


Water treatment

- Real cleanrooms are shown in the next two pictures (from the Siemens compound in München-Perlach)



- The building on the lower left are cleanrooms; the "little" one (**1000 m²**) to the left (with the blue topping) was the **1 μm** research line, the two bigger ones (at right angles; **2000 m²** each) were used for the development of the **4Mbit** and **16 Mbit DRAM** and for the pilot production.
- Below an enlargement



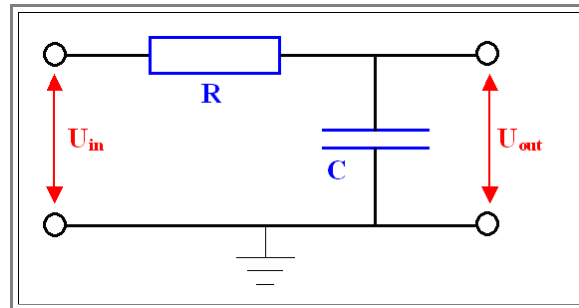
- The yellow emergency exits indicate the actual cleanroom. There are several stories above, and two stories - not visible of course - below.
- The connecting building houses parts of the common infrastructure:
 - Air intake and initial processing
 - Water plant
 - Recycling and cleaning of liquids
 - Cleanroom control
 - Main entrance for heavy equipment
 - Shipping and receiving

RC Time Constant

Basics

This will be a short statement of very elementary facts from basic physics or electrical engineering.

- Any "wire" used to carry electricity has a resistance R (we will not consider superconductors here) - just measure it via $R = U/I$. Its resistance measured in Ohms (O) might be small, but it is never zero.
- Any "wire" that serves its purpose is isolated from the environment by some dielectric (air or some insulation material) having some dielectric constant ϵ_r
- At some distance from any wire is always another conductor at some arbitrary potential. Even if our hypothetical device consist of a wire only, there is always the "earth" somewhere at (by definition) zero potential. It follows: Any wire has some capacitance C to something else. Again, this unintentional **parasitic capacity** might be small, but it is never zero. Moreover, whatever value it has, it is proportional to the dielectric constant ϵ_r of the dielectric in question.
- Any wire (at not too high frequencies) thus can be described by an equivalent circuit diagram that looks like this.



What happens if we put a digital signal on one end of the wire, which we call the "input" end? Ideally, that means that the voltage or better potential on the wire (the not grounded upper "wire" in the picture above) jumps from **0V** to, say, **5V** instantaneously. What happens at the output?

- Ideally, the potential would also go up suddenly after a certain time t_0 which is dictated by the speed of light c , because nothing can move faster than that. we thus have $t_0 = l/c$ with l = length of the wire, and c = speed of light "in" the wire, whatever that means.
- However, as soon as we raise the potential at the input, we have to put charge in the parasitic capacitor C . The voltage at this capacitor is the output voltage, and for a voltage U we need to have the charge $Q = C \cdot U$ stored in the capacitor.
- For that a current has to flow into the capacitor, and that current will be restricted by the resistance R . Note that for just transmitting information as jumps in the potential, current flow would not really be necessary, but our parasitic capacitor needs charges flowing in (and later out) of it, if its potential is to change.

It is easy to describe quantitatively what happens. For the current I flowing during the charging of the capacitor we have two equations.

$$I = C \cdot \frac{dU_{out}}{dt} = \frac{U_{in} - U_{out}}{R}$$

- This is an easy differential equation. We will just look at the solutions for the two cases of U_{in} going "up" to U_{in-on} at $t = 0$ and at U_{in} going "down from " to U_{in-on} to **0**; again at $t = 0$.
- For the "going up" case we obtain

$$U_{out} = U_{in-on} \cdot \{1 - \exp(-t/RC)\}$$

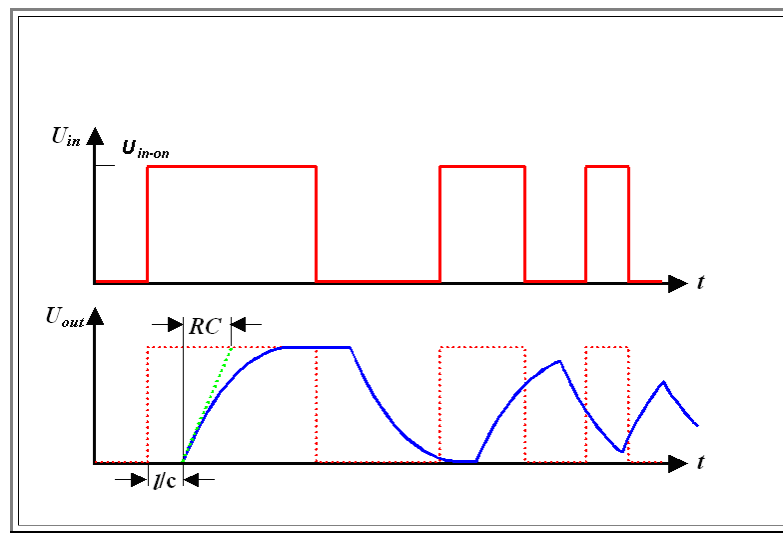
- This describes an output voltage that increases as a function of time with the **time constant** RC ; i.e. after the time RC the output voltage is $U_{out} = U_{in-on} \cdot (1 - 1/e) = 0.63 \cdot U_{in}$.

If we look at what happens if the input voltage is "going down", i.e. switched to zero (after it has been on long enough to make sure the output voltage is equal to the input voltage), we obtain in the same way:

$$U_{out} = U_{in-on} \cdot \exp(-t/RC)$$

- The output voltage thus decays with the time constant RC from the "on" value to zero.

Putting all of this together for some input signals gives the following picture:



What we have now is that at the output end of a "wire" with length l , the voltage will start to go up after a time $t_0 = l/c$ with c = speed of light, i.e. the speed of electromagnetic wave propagation in the wire (c might not be exactly the speed of light in vacuum, but that is not the important part here).

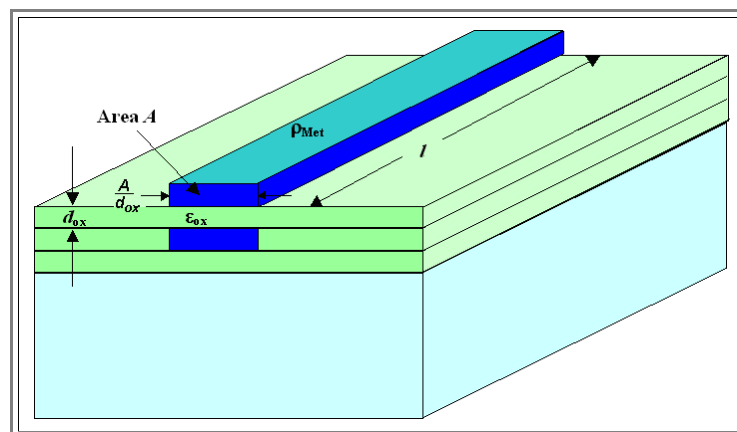
- The voltage then will start to increase, after the **time constant RC** it will be at about **2/3** of the input voltage. After a few time constant it is identical to the input voltage.
- If the input signal goes "down" again, the output signal will follow, again with a delay of t_0 and then with a decay time again given by the time constant RC .

Now consider that we are discussing digital applications. The output voltage will be interpreted as either "0" or "1"; let's say the low potential = 0, the high potential = 1.

- During the transients of the voltage, we are neither here nor there, but our electronic circuitry typically will interpret a potential $< U/2$ as 0, $> U/2$ as 1. That simply means that the time it takes for a signal to travel through a wire with the length l and then to be recognized by what ever follows as a "1" is t_0 plus about one time constant RC of the wire. That means we have an additional delay time given by the "undesirable" properties of the wire.
- Even worse! The picture makes clear, that as soon as you start to transmit signals shorter than a few time constant RC , your output signal can't even follow the input any more. In other words, at input frequencies much larger than about $1/RC$, you run into big problems.

In plain words: You cannot run your circuitry at frequencies $> 1/RC$ - give or take a factor of 2 - 3.!

Where does that leave us? How large is RC typically for the "wires" in an integrated circuit. Let's look at a simplified situation:



Let's look at a typical **Al** or **Cu** conductor on a chip with a length of **1 cm**, a cross-sectional area of **$0.5 \mu\text{m}^2$** and the specific resistivity of [a typical metal of about \$2 \mu\Omega\text{cm}\$](#) . The total resistance of our "wire" then is

$$R = \frac{\rho \cdot l}{A} = \frac{2 \cdot 10^{-6} \Omega\text{cm}^2}{0.5 \cdot 10^{-8} \text{cm}^2} = 400 \Omega$$

- The capacitor in the picture is formed by the two longish "plates", each with an area $l \cdot w = l \cdot A/d_{ox}$ (w is the lateral extension of the conductor strip; i.e $w = A/d_{ox}$). With typical numbers like $d_{ox} = 300 \text{ nm}$, dielectric constant $\epsilon = 3.7$ (the value for **SiO₂**), the capacity of the wires then is

$$C = \epsilon \epsilon_0 \frac{A \cdot l}{d^2_{\text{ox}}} = \frac{3.7 \cdot 8.85 \cdot 10^{-12} \text{ A} \cdot \text{s} \cdot 0.5 \cdot 10^{-12} \text{ m}^2 \cdot 10^{-2} \text{ m}}{9 \cdot 10^4 \cdot 10^{-18} \text{ V} \cdot \text{m} \cdot \text{m}^2} = 1.82 \cdot 10^{-12} \frac{\text{A} \cdot \text{s}}{\text{V}} = 1.82 \text{ pF}$$

For our time constant **RC** we obtain

$$R \cdot C = 400 \cdot 1.82 \cdot 10^{-12} \frac{\text{A} \cdot \text{s} \cdot \text{V}}{\text{V} \cdot \text{A}} = 7.28 \cdot 10^{-10} \text{ s}$$

The maximal frequency we can transmit through this "wire " then would be $\nu_{\text{max}} = 1/RC = 1.36 \text{ GHz}$. This may appear a bit "handwaving" or just a rough estimation; nevertheless, the problem should be clear.

- On a real chip, signals travel through many wires; possibly far shorter then **1 cm**, and possibly not all the way atop another wire, so the capacity could be somewhat smaller. On the other hand, on a real chip, there are also neighboring wires to the left and the right, which increases the capacity.
- Nowadays, we always use a multilevel metalization scheme and being smart, we will keep the small wires on the lower levels quite short, and give the long wires (i.e. for power supply) on the upper levels a large cross-section - keeping **R** down in both cases. Nevertheless, a clock frequency of **4 GHz**, standard nowadays on many mass-produced microprocessors, is an amazing feat considering the number from above.
- Essentially, the specific resistivity ρ of the conductor and the dielectric constant ϵ of the "**intermetal dielectric**" limit the maximal frequency of the chip after all tricks of an optimized geometry have been exhausted.

Decreasing ρ by a factor of about **1.6** by switching from **Al** to **Cu** started to make a lot of sense around the year **2000**.

- Replacing the ubiquitous **SiO₂** with an $\epsilon \approx 3.7$ by a so-called "**low k**" dielectric with an ϵ of **1.5** or so would make a lot of sense right now (**2005**), unfortunately, nobody knows exactly how to do it (despite a Billion \$ or so, that have been spent on the search for a low-k material up to now).

Multiple Choice Test zu

5.0.1 Required Reading Transistors

Start Multiple Choice

Multiple Choice Test zu

5.1.1 What is Integration?

Start Multiple Choice

Multiple Choice Test zu

5.1.2 Integrating Transistors

Start Multiple Choice

Multiple Choice Test zu

5.1.3 Integrating Connections

Start Multiple Choice

Multiple Choice Test zu

5.1.4 Integrated MOS Transistors

Start Multiple Choice

Multiple Choice Test zu

5.1.5 Integrated CMOS

Start Multiple Choice

Multiple Choice Test zu

5.1.6 Basic Considerations for Process Integration - Summary

Start Multiple Choice

Multiple Choice Test zu

5.2.1 Chips on Wafers

Start Multiple Choice

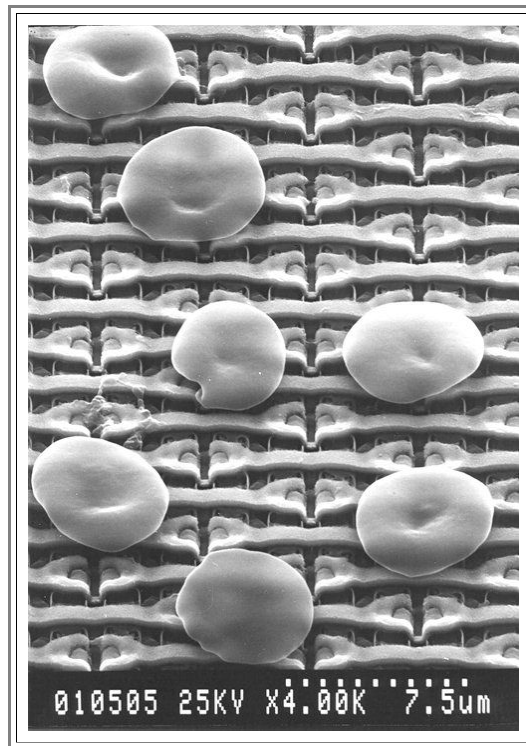
Multiple Choice Test zu

5.5.1 General Aspects of Si Technology - Summary

Start Multiple Choice

Particles on Chips: Blood Cells

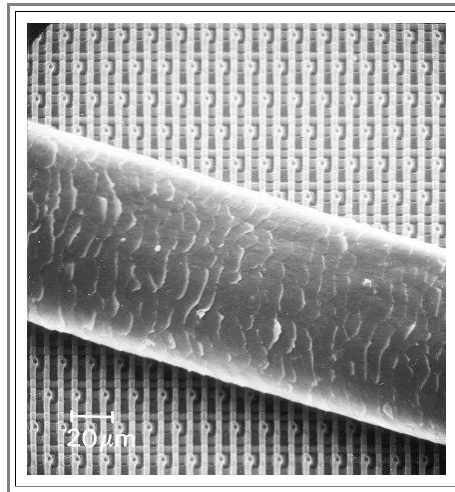
Here is the large size picture of red blood cells on a **1 Mbit** memory



Illustration

Particles on Chips: Hair

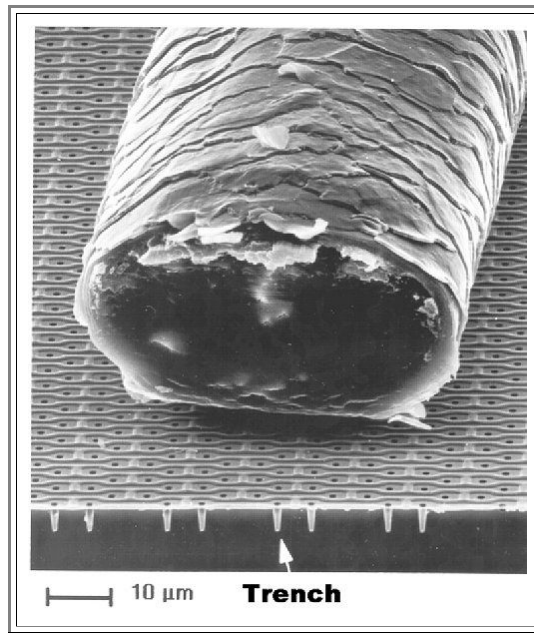
Here is the large size picture of a hair (from a female) on a **256 kbit** memory chip. Just one of the little flakes of the hair would be enough to cover one memory cell.



Illustration

Particles on Chips: More Hair

Here is another the large size picture of a hair on a **4Mbit** memory chip. It also show hows structure sizes decreased - [compare with the hair](#) on a **256 kbit DRAM**.

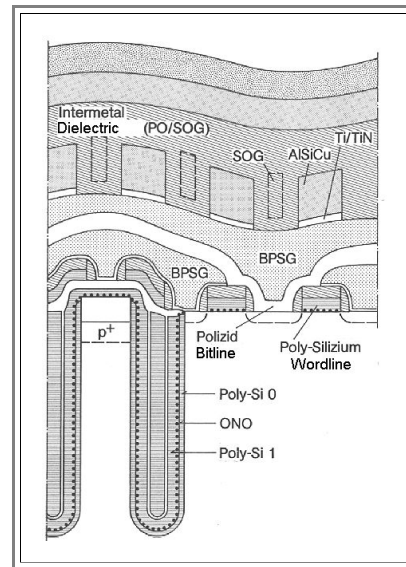
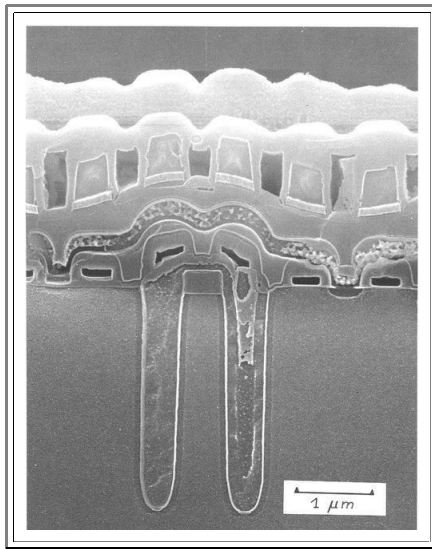


Illustration

Cross Section of 16 Mbit DRAM and 64 Mbit DRAM

Illustration

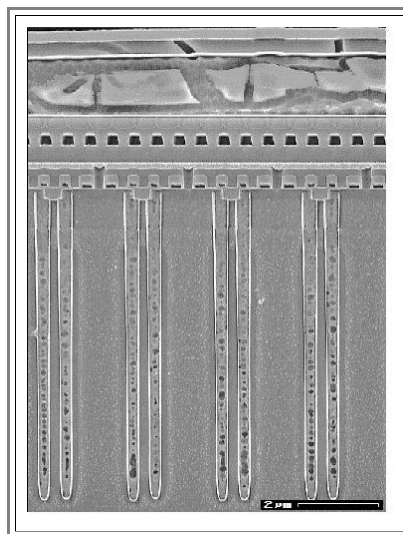
Below a large **SEM** micrograph showing the cross-section of an (early) **16 Mbit DRAM**.



Some explanations:

- The two deep "trenches" (they are really holes) contain the capacitors. Their dielectric (with ca. **7 nm** far too thin to be visible) is "**ONO**", a triple layer of Oxide - Nitride - Oxide.
- The trench is lined with poly-**Si** as a first electrode and as the second electrode.
- To the left and right two transistor gates are visible. The sources of both transistors is the (poly-**Si**) electrode lining the trench and the diffused areas being contacted by the "Polysil bitline". "Polysil" means a double layer of poly-**Si** and **MoSi₂** Molybdenum-silicide.
- The "poly-**Si** wordline" runs perpendicular to the picture and connects the gates of the transistors.
- The "**BPSG**" layers denote **SiO₂** doped with **B** and **P** that serves as insulating dielectric. It is essentially a glass.
- Parallel to the word lines are **Ti/TiN/AlSiCu** lines. They contact the wordlines every once in a while to decrease the ohmic resistance. They consist of a layer sequence: **Ti**, **TiN**, and **Al** doped with about **0,5%** of **Si** and **Cu**.
- On top of this first metal layer is another one running across the picture.
- The metals are insulated by the intermetal dielectric composed of plasma-oxide (**PO**) that contains spin-on-glass (**SOG**) in the interstices.

Below the successor of the **16 Mbit DRAM**, the **64 Mbit DRAM** from a development stage around about **1996**.

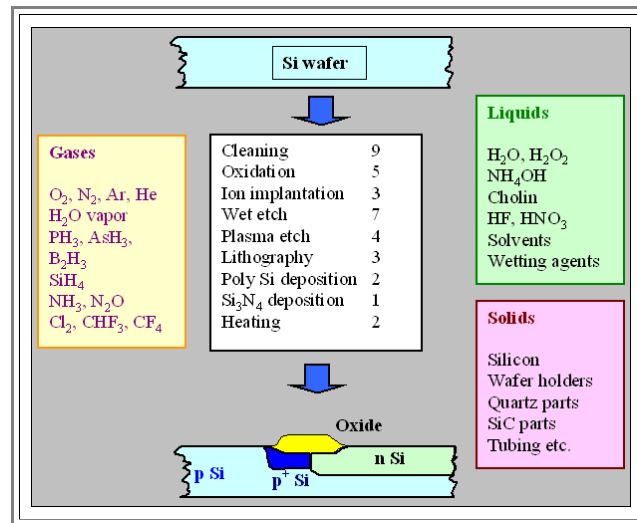


- The structure is essentially the same, but all layers have been planarized.

Processes and Materials I

Here is the list of processes and materials needed for the **16 Mbit DRAM** (an about **1999**) in a graphic way.

Illustration



Note that any material that comes in contact with the wafer or with materials that will come into contact with the wafer, is an **electronic material** - utmost care has to be taken in selecting the right stuff!

- Wafer holders or tweezers, e.g., can totally ruin a wafer by leaving minute amounts (far below the detection limit) of heavy metals (most notably **Fe**) on the wafer if they are unsuitable (Never, really never, touch the wafer with a metal tweezers!).
- Gas pipes may corrode internally if made from the wrong metal and thus contaminate the gas flowing through it with traces of impurities - your factory then will only produce garbage.

Note also that some of the most dangerous inorganic chemical are used!

- HF** (hydrofluoric acid) will cause heavy tissue and especially bone damage already by its vapors - you do not even have to touch it to get severely damaged.
- PH₃** (phosphine) and **AsH₃** (arsine) are among the most toxic gases known to mankind; minute amounts are deadly (PH₃, in fact, was used as a poison gas in world war I).

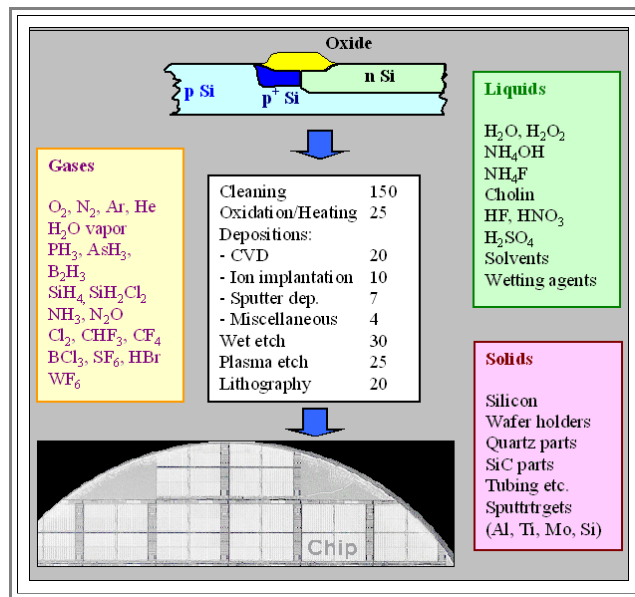
To continue, use the link

Processes and Materials II

Here is the rest of the processes

- A few more materials are needed, especially solids in the form of "sputter targets".
- A great total of about **450** process- and control steps are needed.

Illustration



We don't have a chip yet - we only have unpackaged chips on a wafer.

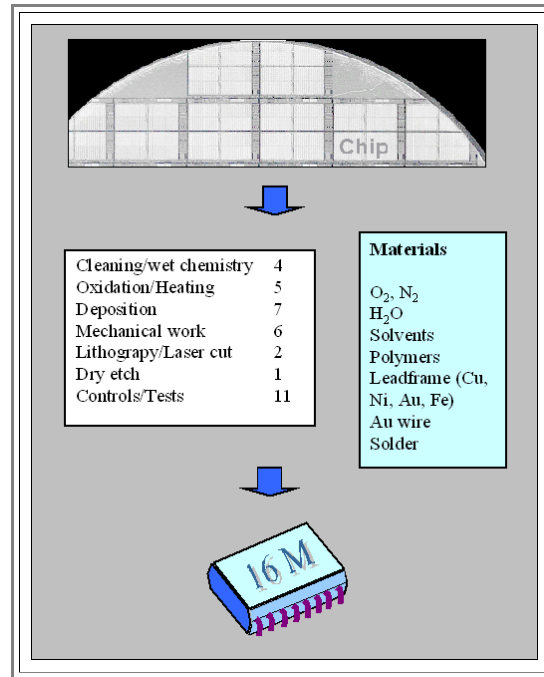
- Next, the wafer has to be cut and the chips that work (this needs a measurement) are packaged.
- Packaging, although not needing processes at very small dimensions, is not simple either.
- To continue, use the link**

Processes and Materials III

■ Packaging is a completed process in its own right

- It needs very special materials - even the lowly black plastic that dominates the appearance of chips is a sophisticated material!
- It also includes extensive testing of the chip.

Illustration



Wafer Flats

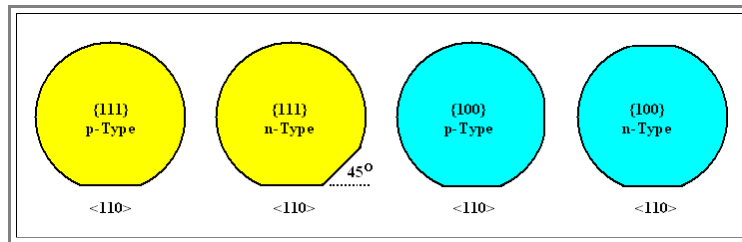
Illustration

In the old times (up to the **150 mm** wafer diameter era), wafers had **flats**, and the flats told you two things:

1. The doping type of the wafer (**n**- or **p**-type)
2. The orientation of the wafer: **{100}** or **{111}**

- While this is trivial information, consider: All wafers, whatever doping type or crystal orientation, look exactly the same! As soon as a wafer has been removed from its box that carried this and other information, you can't see anymore what you got. You also cannot measure it easily (and without destroying the wafer).
- And if something goes wrong here (and things that can go wrong will go wrong some day), it may be a horribly expensive mistake! If you feed wafers of the wrong doping kind into the line, it will really, really cost you - probably your job.

So here is the convention



But beware! Wafer manufacturers will produce whatever the customer wanted, and after **n**-type **Si** went out of style for most mass-produced chips, the only reason for a flat was to allow the patterns to be made to be aligned with a crystallographic direction.

- You then ordered your **p**-type **{100}** wafers with only one flat in the **<110>** direction. One reason was that the wafer would easily cleave along this and the respective perpendicular direction.
- So wafers with diameters larger or equal to, say, **100 mm** and just **one** flat are more likely **{100}** **p**-type than the "proper" **{111}** **p**-type. The [picture in the backbone](#), e.g., shows **p**-type **{100}** **150 mm** wafer!
- And wafers with diameters larger or equal to, say, **200 mm**, probably will have no flat at all, but just a small "**notch**" - simply because you lose too much expensive area by cutting of a flat.

So, how can you tell what you have - if you don't trust the one flat there is, or if there is none! There are extremely simple ways of checking:

- Checking doping type:** Take an Voltmeter and measure the **thermoelectric voltage** between a hot tip (take a soldering iron) and a room-temperature tip pressed on the wafer somewhere. Its **sign** will tell you if the wafer is **n**- or **p**-type. Which is which results from hard thinking or from checking a known piece of **Si**.
- Checking orientation:** Break your wafer. If the pieces tend to be rectangular, it was **{100}**.

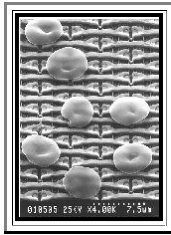
In either case your wafer is now "dead", i.e. no longer usable for making **IC's**.

Particles on Chips

Here you can see a collection of **SEM** pictures which not only illustrate graphically the "particle" problem in making chips, but also have a certain esthetical appeal

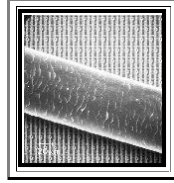
Click on the link for an enlarged view.

Illustration



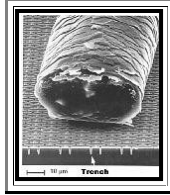
Red blood cells on a **1 Mbit** memory

[\(large size\)](#)



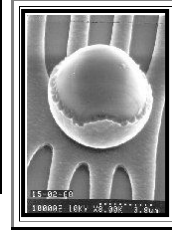
Hair (from a female) on a **256 kbit** memory chip. Just one of the little flakes of the hair would be enough to cover one memory cell.

[\(large size\)](#)



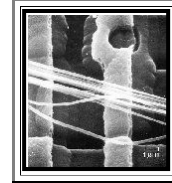
Hair on a **4Mbit** memory chip. - just to show how structure sizes decrease.

[\(large size\)](#)



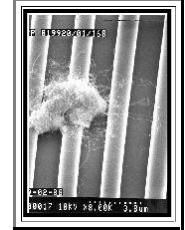
Here a metal particle, probably a tiny drop of **Al** that was burnt off by a electrical discharge in a sputtering machine and hit the **Si** as a solidified droplet. It was coated with **Al** which was subsequently structured by etching. Four conducting lines are now short circuited.

[\(large size\)](#)



Mother nature is still ahead when it comes to small structures. Here we see spiderwebs on a **256 k DRAM**. A typical strand of spider silk consists of several individual strings with diameters around **0.2 μm**.

[\(large size\)](#)



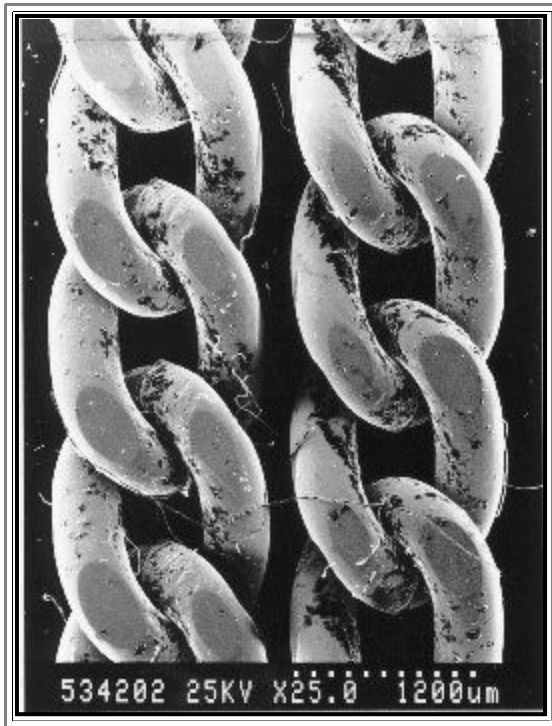
Here is a particle of unknown nature. Whatever it is, it will kill a chip.

[\(large size\)](#)

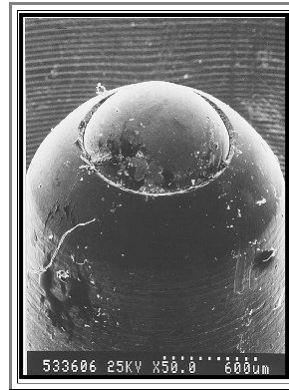
Humans as Source of Particles

Here a few examples of humans as source of particles. The magnification is typically low; even particles not visible under these conditions would be deadly

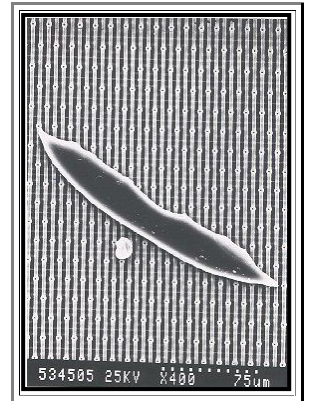
Illustration



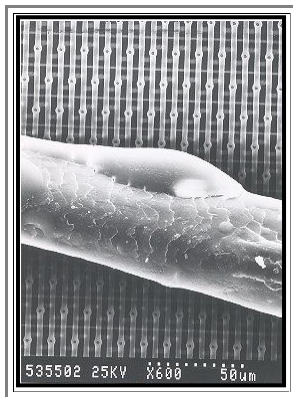
Dirt on a necklace



Tip of a ball point writer



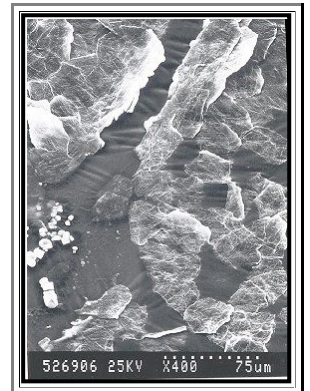
A piece of nail polish on
a **256 k DRAM** chip



Hair with hair spray



Inside of finger ring



Flakes of skin

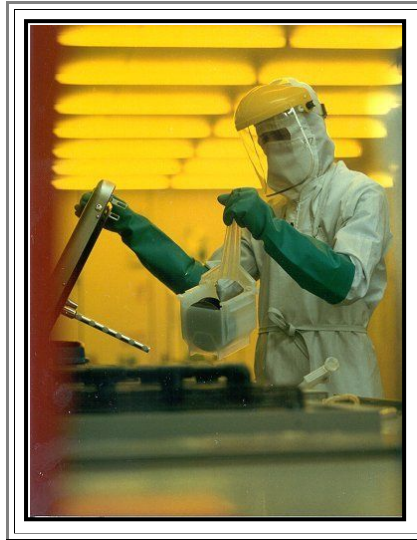
Cleanroom Garments

Below some pictures from inside a cleanroom. We also see some pieces of equipment.

Illustration



General view of work in a clean room



Inserting wafers into a "cleaner"; a kind of washing machine for wafers employing extremely aggressive chemicals. Special protection (heavy gloves and a face shield) are worn.



Looking along the mechanism for moving wafers into a furnace. The orange glow from the furnace tubes is visible.


Electronic Materials in Unexpected Places

Below an add (from "Semiconductor International"; June **2000**) that shows that nothing is too unimportant to qualify as electronic materials,

Illustration

Contamination got you floored?

Contec® presents the Edgeless Mopping System

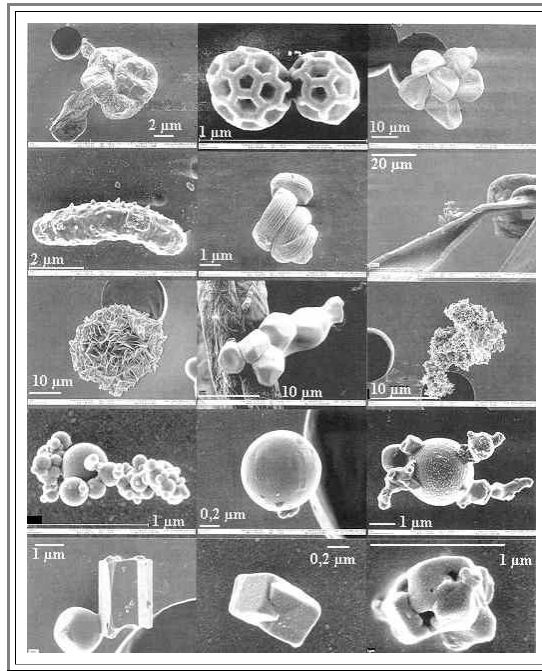


Edgeless Mopping System

- Cleanest floor mop available for cleanroom use
- Continuous tube of 100% polyester knitted fabric looped to form tubular mop strands
- Strands encapsulate liquids to maximize sorptive capacity
- Eliminates fiber generation during mopping
- Specifically engineered for wet mopping cleanrooms
- Laundered and packaged in a Class 10 cleanroom
- Mop is available sterile
- Fully autoclavable
- Stainless steel or ultraplasic handles, bucket, wringer and basket insert complete the system

The Air We Breath

Here is a picture of the actual air we breath; it was assembled by the Max-Planck-Institut für Chemie in Mainz/ Germany



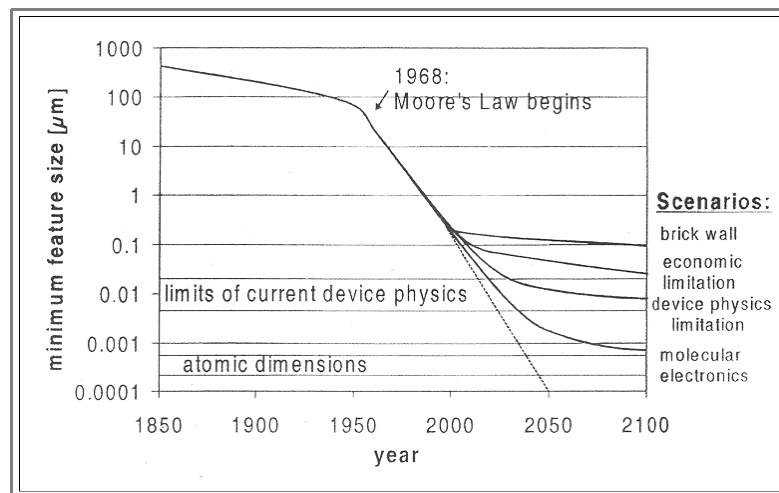
What we have are

- Biological objects (pollen and such) in the first two rows and middle of the third row.
- A silica particle (from the desert?); left third row; and an unknown particle right, third row.
- Ash particles in the fourth row.
- anorganic salt crystals (we think) in the fifth row .

Moore's Law: Break Down Scenarios

Illustration

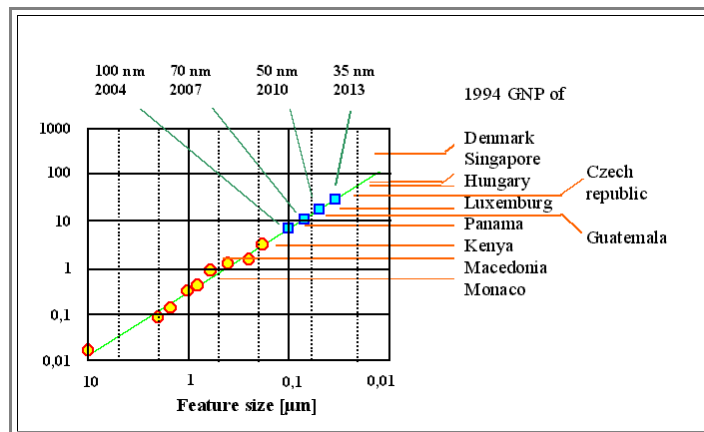
Here is an illustration of Moore's law with several scenarios for possible extrapolations.



The dashed line shows the linear extrapolation. It is sure to break down - at the very last when feature size come down to the level of atoms

- The "molecular electronics" scenario thus must be seen as the best possible case and would keep us going for another **20 - 30** years.
- On the other extreme, the "Brick wall" scenario assumes that technology simply hits an insurmountable barrier right now. In other words, feature sizes will not come below about **0,1 μm** ever. The end then would be near.
- In between are scenarios where feature size reduction is either limited by the costs (in other word, you can make it, but it is to expensive to sell) and device physic limitations. After all, a conventional **MOS** transistor needs at least a few doping atoms in its source and drain region and since the density of doping atoms must be considerably smaller than the density of **Si** atoms, you cannot make the transistor components arbitrarily small.

What "economic limitations" mean is illustrated in the next picture

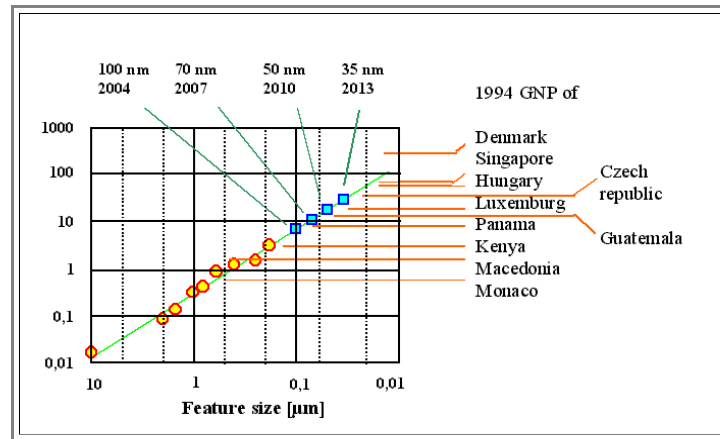


- Here the costs of chip development is plotted (circles) and from the extrapolation of the resulting rather straight line predicted for the coming years (squares) and compared to the gross national product (**GNP**) of sizeable countries
- The problem will be to recover these costs at roughly the same prize per chip. It means that you have to sell at lot more chips at each generation. Assuming that the capacity in memory chips increases fourfold every three years, the market for bits then has to grow much faster than about **60%** a year!

Costs of Chip Production

Illustration

This graph shows the development of the costs for developing a new chip generation (including the investments needed for building the factory) in comparison to the gross national product (GNP) of some nations.



There is an exponential increase in cost!

- While a **1 Mbit DRAM** (**1,0 µm** technology, about **1990** vintage) could be had for as little as **\$ 2.000.000.-**, the **1 Gbit DRAM** (**0,1 µm** technology, available in about **2003**) will run up a bill of about **\$ 7.000.000.000.-**
- Compare this to the gross national product of some countries and you will see that "globalization" is more than a catch word for high-tech products!

Here the latest news that fits right in with the table above

- From **Oct. 2002**

Intel Opens \$2B New Mexico Fab

Online staff -- 10/23/2002 **Semiconductor International**
Electronic News

Intel Corp. today announced the opening of a **\$2 billion dollar** expansion to its manufacturing facility in Rio Rancho, N.M., adding 200,000 square feet of clean room space.

Designated Fab 11X, the plant will produce microprocessors on 300mm wafers using 0.13-micron process technology and will transition to 90nm process technology in 2003, Intel said.

"As computing and communications devices converge, the need for increasingly complex components with more capabilities will grow," said Paul Otellini, Intel's president and COO. "This facility will help us meet that growing demand. The combination of the 300mm wafers and 90-nanometer process technology will also reduce the costs of manufacturing, increase productivity and improve the availability of the world's most advanced semiconductor products."

In addition, Intel said that from an environmental perspective, water and chemical use will be more efficient at the plant, claiming that when compared to a 200mm facility, Fab 11X will produce 48 percent less volatile organic compound emissions, use 42 percent less ultra pure water and will use about 40 percent less energy.

- From **Oct. 2003**

Semiconductor International, Oct. 2003

It seems that hardly a day passes without a new announcement about the cost or economics of 300 mm wafer fabs: "\$7.0 billion dollars estimated revenue to support a 300mm fab,"¹ "**Leading-edge fab costs soar to \$4 billion**,"² "Prices for 300 mm wafers remain too high, says analyst,"³ "Revamping 200mm fabs will have a significant cost advantage over new 300mm facilities for some products for some time."⁴

- From **July 2005**

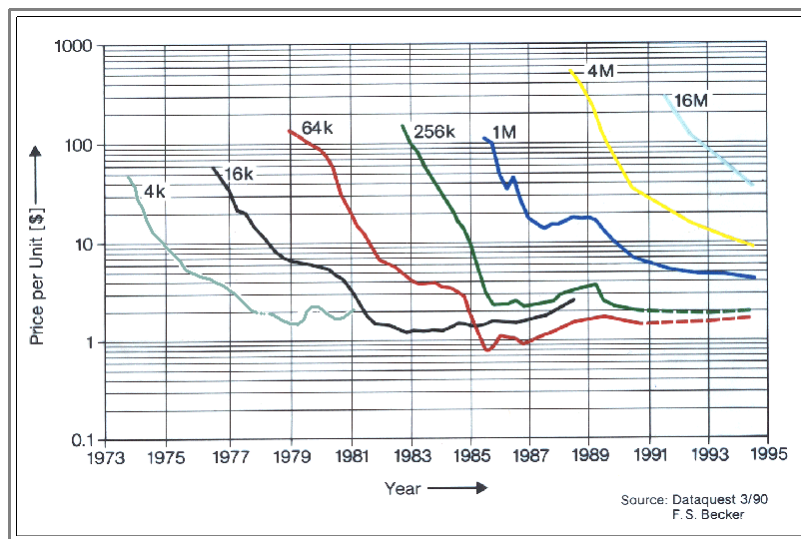
Solid State Technology, Aug. 2005

Intel Corp., Santa Clara, CA, announced on July 25 that it will build its sixth and newest 300mm/45nm fab at the site of its Chandler, AZ operation. Construction will begin immediately on the **\$3.0 billion**, 1.0 million sq. ft Fab 32 facility, with production slated for 2H07. Intel operates four 300mm fabs in the US and Ireland, with a fifth expected to begin operations later this year -- a \$2 billion conversion of an existing 200mm fab, also in Chandler.

Prize Decay for Memory Chips

Illustration

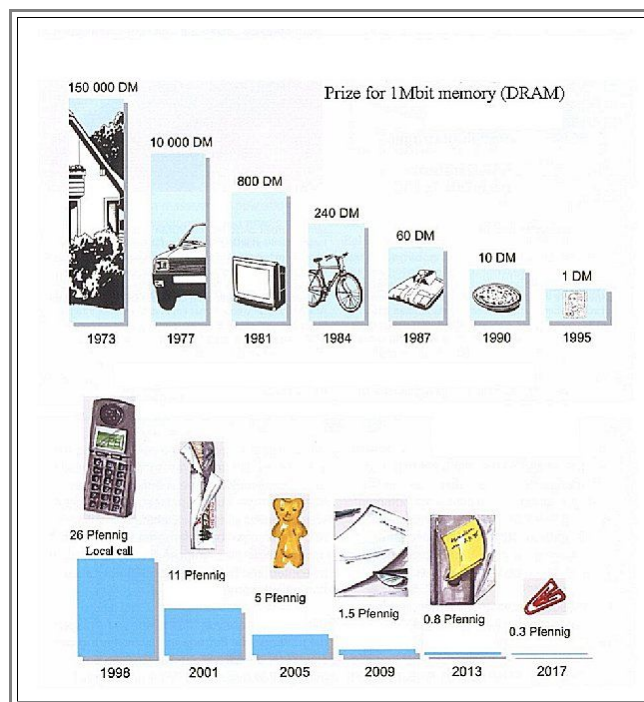
Shown are the average prizes for **DRAMs** as a function of time for each generation



Note that the scale is logarithmic!

- While you may get around **\$100** for chip for a short time, you will be down to a few \$ within **3 - 4** years!
- There are influences beyond your control. In bad years the Dollar exchange rate may kill you (chips are always sold in \$). If you are lucky, trade restrictions (by the Americans in **87/88**) make the commodity scarce and prizes go actually up (see the **1 Mbit**; very unusual!).

What this prize decay really means is very graphically illustrated below where the cost of **1 Mbit** of 1 Mbit memory is expressed in every day items

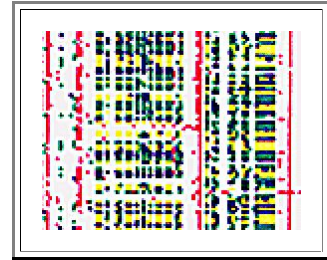
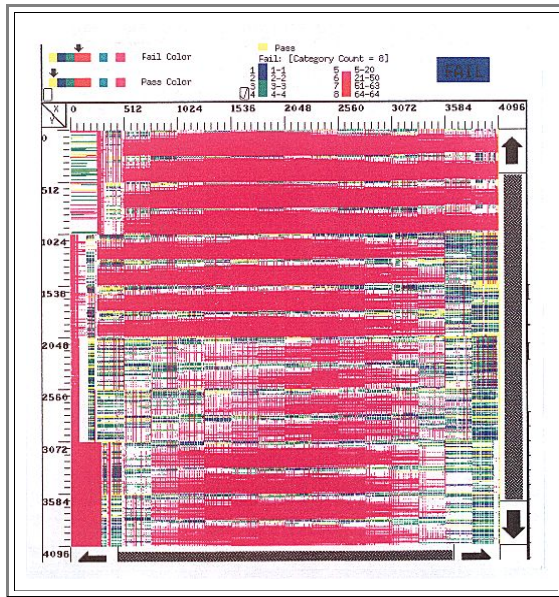


Emergence of a New Chip

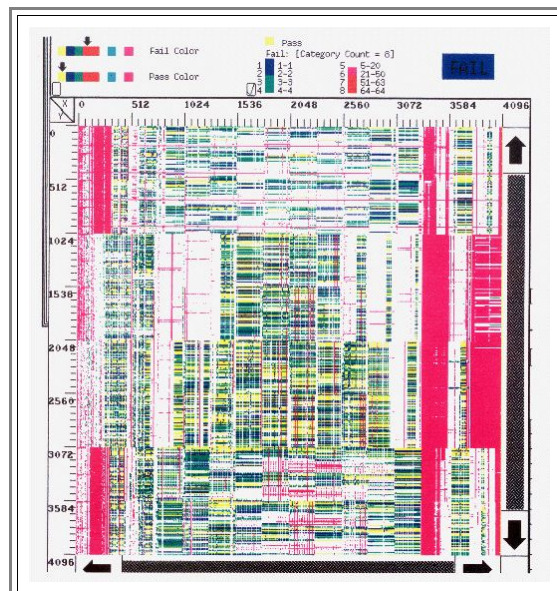
Illustration

After much designing, simulation, process trials and test chips comes the moment when the first product chip containing the new generation will enter the line and becomes processed. Processing may take several month because many processes do not work too well or are done in prototype machinery. Here is a real example from the development of a **16 Mbit DRAM** at Siemens.

- If the batch of wafers does not die a sudden death at some shaky process (or because somebody simply drops them), "**First Silicon**" will eventually emerge. This is a big moment in the life of the team - now we will see if several **100** man years and many million **\$\$** were well spent.
- The test crew starts to measure the function of the memory with the chips still on the wafer. The results are plotted in a matrix mimicking the matrix of memory cells on the chip. The color given to a pixel encodes the quality of the individual memory cell. Yellow means fully functional, all other colors, especially red, mean trouble.
- Below is a plot for the best chip from the "first Silicon". The insert on the right shows that single pixels are resolved in principle. How was the feeling with a result like that?

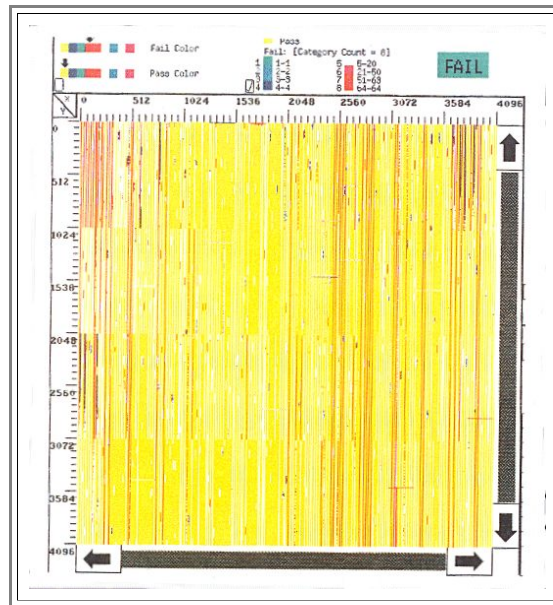


- The feeling was : *Get out the champagne!* There are fully functional memory cells on this chip! Not too many - but who cares. This was good! More often than not, nothing whatsoever works on first Silicon. A few hundred cells that work are enough to measure all critical parameters of the circuits and to get the desperately needed data for optimizing design and processes.
- Next comes *second Silicon*. Those wafers probably were processed to some degree when first Silicon came out, so only back end processes can now be optimized. But there is improvement: A little bit more yellow, but far less deep red!
- The structure in these plots - pixels of the same color arranged in lines - simply reflects the fact that a number of memory cells share the same bit- or word line. If this line is interrupted, all its memory cells fail.

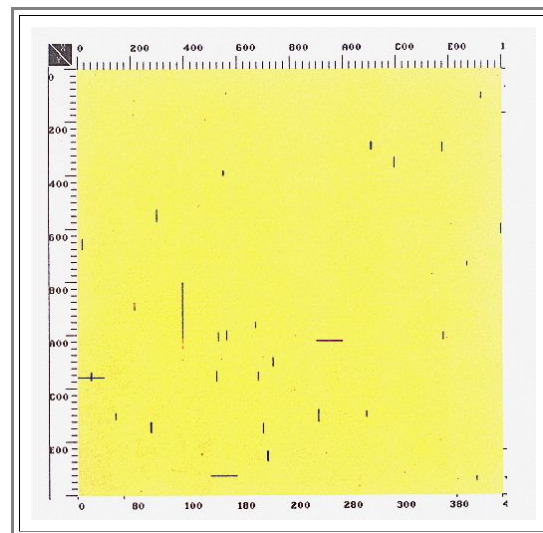


Here we have the first redesign - all the things learned from first Silicon are now incorporated into the chip.

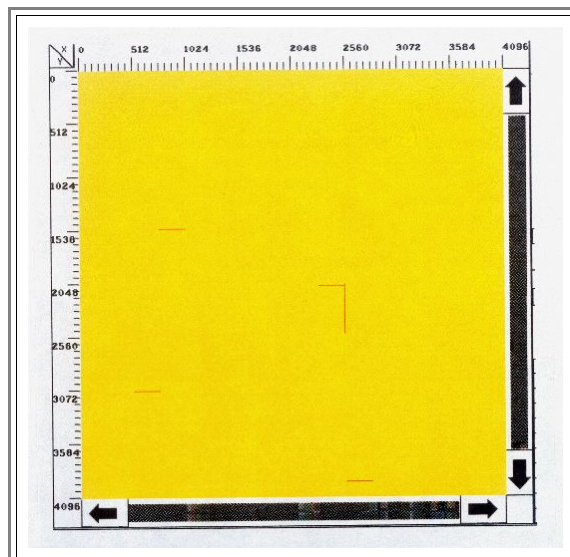
- And now the yellow color dominates. While this looks pretty good, we are still a million or so cells short of total success.



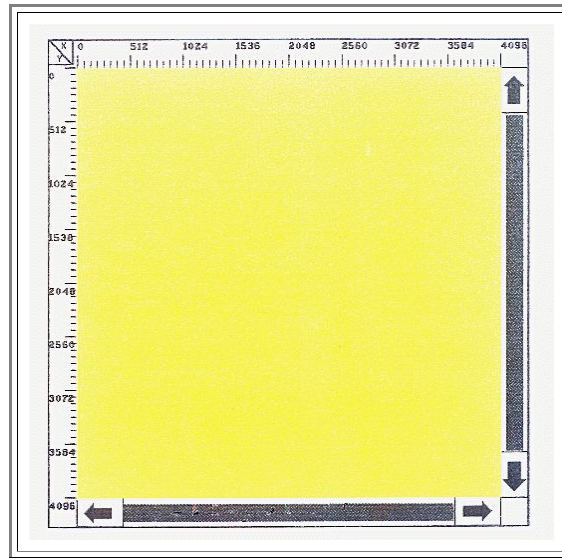
Now the process stabilizes and a new game starts: Get rid of the few but noticeable problems in the memory matrix. This is tough!



Now we won: Only a few strongly correlated failures - no single bit problems. This can be solved - with luck - by activating some redundant memory cells that will take over the function of the failed cells



Triumph!! Using the built in redundancy, a fully functional memory matrix was obtained for the first time - in the third full process cycle, a very good score.



Exactly two chips out of some thousand candidates are fully functional. The project goals have been met.

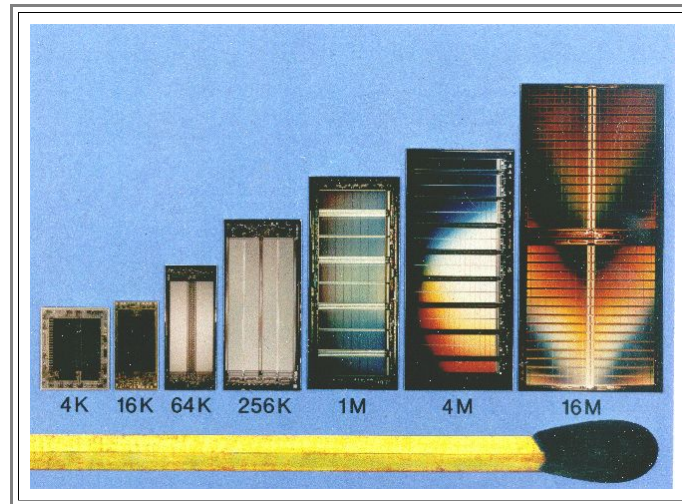
All that remains to be done now is to crank up the yield to, say, **15 %** functioning chips per wafer so that production can start.

● *That was a joke:* Now the really *hard* work starts. It will take considerably more time and money to "just improve the yield" than it took to get to this point in the chip development.

Development of the Chip Size

Here is a very graphical picture of the development of the chip size for **DRAMs** up to the beginning of the nineties.

- You must be aware, however, that chip sizes are not static. As soon as your first product is finished, you try like crazy to reduce the chip size.
- That is almost the only way to make money: Make *more* chips with the established (already paid for) technology.

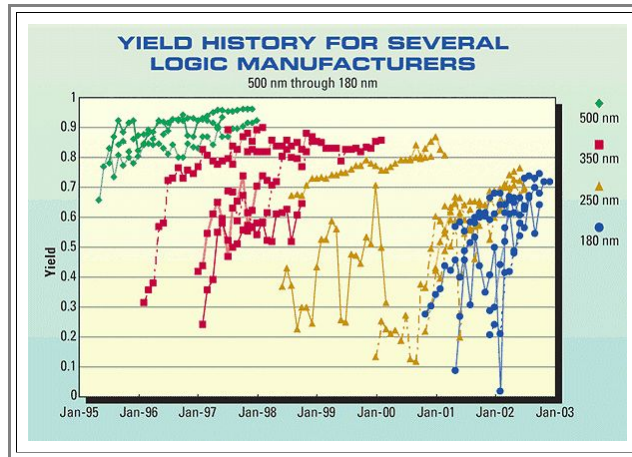


Yield Development

Illustration

In the backbone part, there is a [hypothetical yield over time curve](#), and some far-reaching conclusions are drawn from it.

- What do *real* yield curves look like? Well, few people know, because that is one of the bigger company secrets.
- However, in **2003**, Leachman and Berglund, working for *Sematech* did a study and release some real (anonymous) data. Here is the interesting curve:



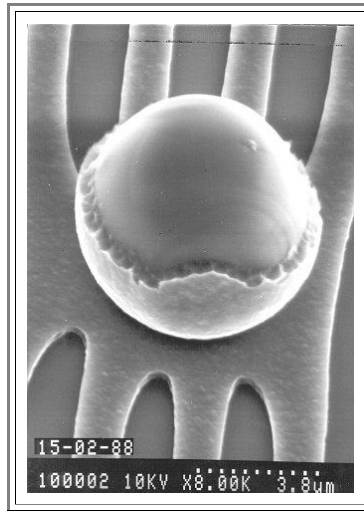
The yield may go up somewhat faster than on the hypothetical curve, but several things are quit clear:

- Companies start production with yields as low as **(10 - 20) %**: In other words: Most of the chips produced go in the garbage bin!
- There is a lot frustration potential: For no recognizable reason your yield suddenly collapses substantially! (If the reason would have been recognizable, it would not have happened for more than a day or two and then not shown up on the monthly data given above!)
- It's rather difficult to have yields in excess of **80 %** - and it is getting worse.

Particles on Chips: A Metal Ball

Here is the large size picture of a metal particle, probably a tiny drop of **Al** that was burnt of by a an electrical discharge in a sputtering machine and hit the **Si** as a solidified droplet . It was coated with **Al** which was subsequently structured by etching. Four conducting lines are now short circuited.

Illustration

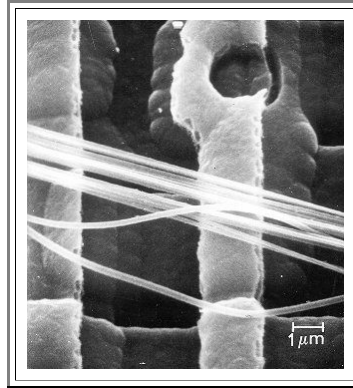


Particles on Chips: Hair

Here is the large size picture of some strands spider silk on a **256 k DRAM**. It shows that Mother Nature was still ahead then (**1988**) with regard to small structures. A typical strand of spider silk consists of several individual strings with diameters around **0.2 μm**

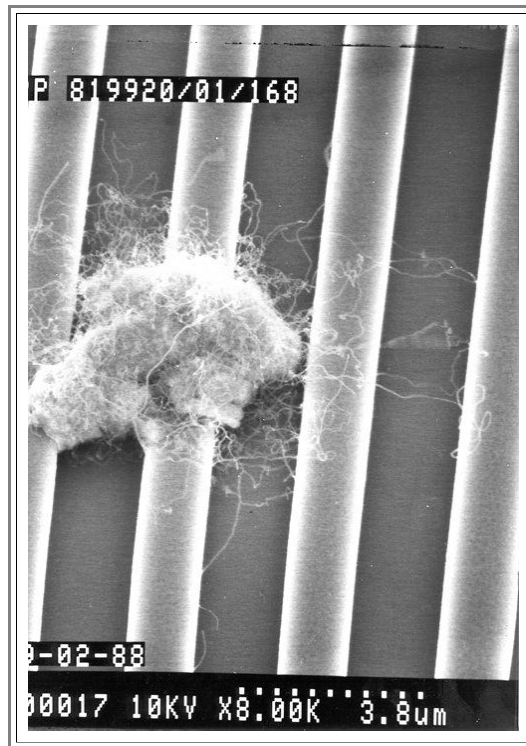
However: Nowadays (**2002**) we are doing better than that. Minimal dimensions on state-of-the-art chips are **0.13 μm** .

Illustration



Particles on Chips: Unknown Object

Here is the large size picture of something unknown (probably biological). Whatever it is, it will kill a chip.

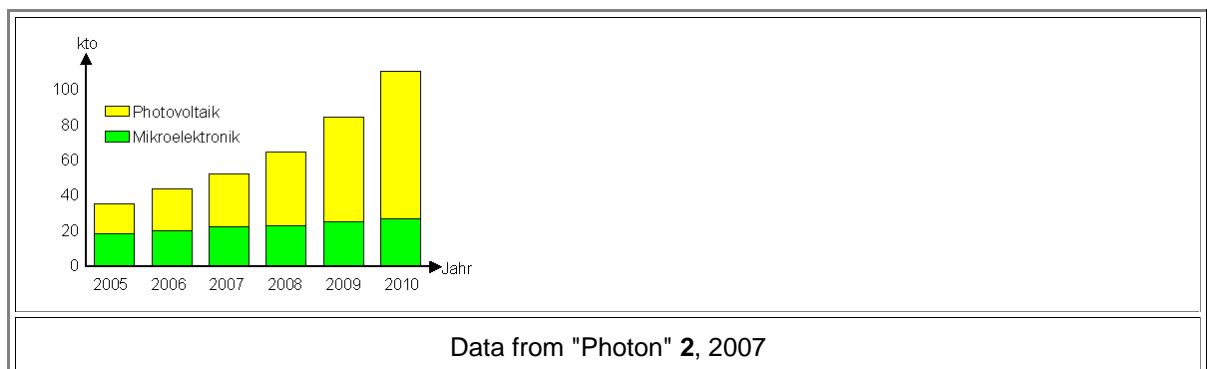


Illustration

The Great Si Crisis in 2007

Advanced

- While today's buzzwords like "energy crisis", "climate change" and "greenhouse effect" have finally penetrated even the most simple minds on this planets (even George W. Bush knows the meaning of this words by now), more sophisticated minds knew about this already in the **70ties** of the last century.
 - That "the future" would only have a future if we would manage in time to switch from burning coal and carbohydrates to energy supplied relying on "solar" - wind, water, biofuel, solar heat and direct solar electricity via solar cells - was clear to everybody who knew a minimum about energy conservation laws and the effect of greenhouse gases. One might consider nuclear power in this context, but we won't.
 - What happened, or better what did not happen after the first and second "oil crisis" in the **70ties** (look it up in the Net), is history now. It can be safely predicted that in years to come when historians look back on this chapter of global history, not everybody who was in possession of power in the second half of the second millennium will look very good
- Anyway, by some cunning politics and good old dumb luck, Germany started a thriving solar industry a few years after the millennium; a few other countries (foremost Japan), did the same.
 - Predictably**, real money was made with the most old-fashioned technology: more or less "classical" solar cells made from bulk **Si** - either from (round) single crystal wafers or from so-called (square-shaped) [multi-crystalline slices](#). Starting around **2003**, factories were built like crazy, and new companies like Q-Cells or Solarwold turned almost over night into stock market stars worth hundreds of Million €
 - Predictably**, [electronic grade Si](#) (even somewhat less "good" **Si** from minor suppliers) turned scarce and prices shot up. Solar cell companies without reliable sources of **Si** (i.e. without long-term contracts with major suppliers) had to give up; while others grew with growth rates only limited by how fast you could hire people and built factories (about **40 %** per year).
 - Given the growth rates of the solar industry **and** of the microelectronic industry and projecting that on the **Si** production capacities of the existing companies (just about a handful), it was pretty clear that electronic grade **Si** would become scarce around **2005**. However, given the general experience in big business, it was also clear that not much would happen **before** this predicted scarcity was really felt.
- Now, in Feb. **2007**, big articles about the **Sicrisis** appear rather regularly in the relevant magazines and lots of money is diverted to electronic grade **Si** production.
 - Here are a few facts; mostly taken from the Feb. **2007** "Photon".
- Major poly-**Si** (always electronic grade now) producers like Wacker (Germany) MEMS (USA) and ... are building new plants for poly-Si production.
 - Besides just cranking up the established production with the ["Siemens" process](#) as described in the backbone; new variants of the basic **CVD** process are being tried and moved to production. Extremely simplified, instead of having a thin **Si** rod grow by depositing **Si** in a relatively small reactor, "**Si** dust" is fed into the top of a huge heated tower, where it slowly sinks down in a strong upward draft provided by silane and hydrogen gas (plus the doping gases) fed into the tower at the bottom.
 - The dust particles grow while sinking; if all goes well, you can eventually shovel out some **Si gravel** at the bottom
 - Sound simple, is not. Just the **PH₃** or **B₃H₆** inside a decent sized reaction tower would be sufficient to wipe out a big city if it ever gets out.....
 - As a very raw but interesting number in this context we note that a production capacity of **10.000 to/a** will cost you **(0.5 - 1) · 10⁹ €** initial investment; you also need the (very secret and closely guarded) know-how.
- Here is a prediction of how much **Si** we will need in years to come. The average growth rate is about **25 % /a**; accelerating as time goes by.



Here some of the present gossip around the **Si** crisis:

- "Orkla ASA" (Norwegian company) builds **323 Mio €** plant for producing solar **Si** by direct cleaning of metallurgical **Si**. Capacity is **5.000 to/a** in **2009**. "ELKEM", a Orkla daughter company, is the world's largest producer of metallurgical **Si**.
- "Dow Corning" (USA company; not much previous **Si** experience) announces end of **2006** (as a surprise) that it will produce **1.000 to/a** of solar-grade **Si** by some process for cleaning metallurgical **Si** (in Brazil).
- 4 other companies are making similar, but less specific announcements.
- About **20** more companies (including Chinese) are suspected to secretly work on some **Si** production process based on cleaning metallurgical grade **Si**.
- **1 kg** of solar **Si** presently costs about **50 €** (if you have a good contract with a supplier), on the "spot market" you may have to shell out **150 €/kg**. Prices have gone up **30 %** and more during the last **3** years.
- The projected **110 kto/a** in **2010** is still not enough for the expected market for solar cells. This may mean, that the thin film solar cells, which so far are just a niche production, may come into there own.

Obviously, the new game in town is to take cheap metallurgical grade **Si** (at around **1 €/kg**) and to clean it "directly", avoiding the avoid costly (**CVD**) Siemens process. It's actually not all that new; I have worked along those general lines around **1991** in the Siemens labs in Munich myself.

- So how is it done? Few people know - whatever is going on is ultra secret, and all the players in this field have their own little "dirty" or better "cleaning" tricks. Before we look at some of those tricks, we will look at the money side of things a bit more.
- The only definition of what defines "**solar grade**" **Si** as a subgroup of "electronic grade **Si**" is how good the solar cell will be that is manufactured with the stuff; in other words what kind of [cell efficiency \$\eta\$](#) (electric power out/ Light power in (in %)) you can get with your process.
- A few rough numbers to that are: If the best efficiency you can get is below $\eta = 13\%$ - forget it. Even if your **Si** would be for free, it's too expensive. On the other hand, if you can get an $\eta = 15\%$, you can ask for about **40 €/kg**. This just shows how dramatically important the cell efficiency is (and how difficult to increase it relative to the best you can do).
- In other words: If you make or loose huge amounts of money by going into the **Si** business depends on many variables most of which you don't know all that well.

Last let's look at some ideas of how to clean metallurgical grade **Si** cheaply and efficiently. First, we have to realize that what works for some impurity many not work for some other, so we have to group impurities in e.g. metals, doping elements, **O** and **C**, and so on. A few typical processes might be

- Melt the stuff and throw some slag-formers in the melt (e.g. **CaSiO₂**) Doping elements like **B** and **P** might prefer to be in the slag, which you then spoon off with a ladle (haha).
- Melt the stuff and blow gases through the melt. With some luck (or special knowledge) - see above).
- Solidify your melt in a suitable way ("directional solidification") and use [segregation](#) to move a lot of the (metallic) impurities into that part of the melt that solidifies last (which you throw out then).
- Bath your (finely crushed) **Si** in suitable acids and lyes and dissolve the dirt sitting on the surface. If you crush material full of defects (particularly grain boundaries) after some treatment that drove your impurities into the defect sites, and if you crush it in such a way that fracture occurs at grain boundaries, you may now chemically remove a lot of the dirt.
- Keep our energy needs down. melting a **kg** of **Si** takes a certain amount of **kWhr**; so does heating your equipment. The Siemens process runs up a bill of about **100 kWhr / kg**; cleaning metallurgical grade **Si** may keep it a **15 kWhr/kg**.
- Don't just take any metallurgical grade **Si** - make your own with "optimized" dirt in there. Find a source of relatively clean coal and quartz - at least with respect to the more trickier impurities.

Let's stop here. If you get my drift, you will now understand that there are many interesting and demanding jobs out there for engineers who know their electronic materials, and that the **Si** crisis can and will be solved very quickly. Exactly how remains to be seen.

- As an afterthought: The situation has many parallels to the [introduction of steel mass production](#) of the second half of the **19th** century. People were fighting impurities, had no clear idea of exactly what was need, and found a lot of working solutions that were only understood later.

Three month after I wrote what's above about cleaning metallurgical grade **Si** we learn from the Semiconductor International Newsbreak, March **26th, 2008**:

Amsterdam, 27 March 2008 --- AMG Advanced Metallurgical Group N.V. (EURONEXT AMSTERDAM: AMG) today announced that its 50.5% owned subsidiary Timminco Limited (TSX: TIM), through its wholly-owned subsidiary, Becancour Silicon Inc. ("BSI") has announced that it has entered into an agreement to supply solar grade silicon to Q-Cells AG. In its press release Timminco stated: Under the terms of the agreement, BSI will supply Q-Cells with contractually fixed supplies of 410 metric tons (mt) in 2008 and 3,000 mt in 2009 at fixed prices. The deliveries start immediately. Until the end of July 2008 the partners will negotiate a further contract on the delivery of up to 6,000 mt per year in the years 2010 to 2013. The price for these possible further supplies will be negotiated contingent upon market conditions. With this contract Q-Cells will become BSI's largest customer for solar grade silicon in 2009.

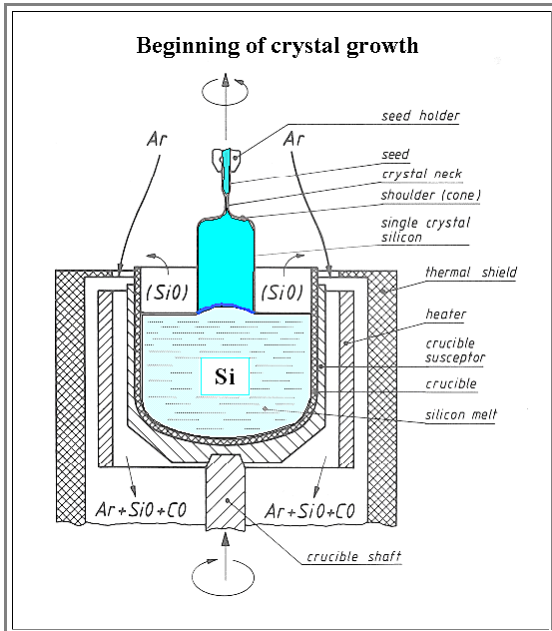
"BSI has developed a proprietary process which enables it to produce solar grade silicon by purifying metallurgical silicon. Q-Cells has tested unblended material from BSI extensively and has obtained very good results in cell production", said Mr. Anton Milner, CEO of Q-Cells. "This contract represents a giant step for our solar grade silicon business. Q-Cells is a leader in the photovoltaic industry and their endorsement of our material through this supply agreement is further evidence of the paradigm shift we are creating in the solar grade silicon market." said Mr. Rene Boisvert, President and CEO of BSI.

The Science and Art of Si Crystal Growth

Advanced

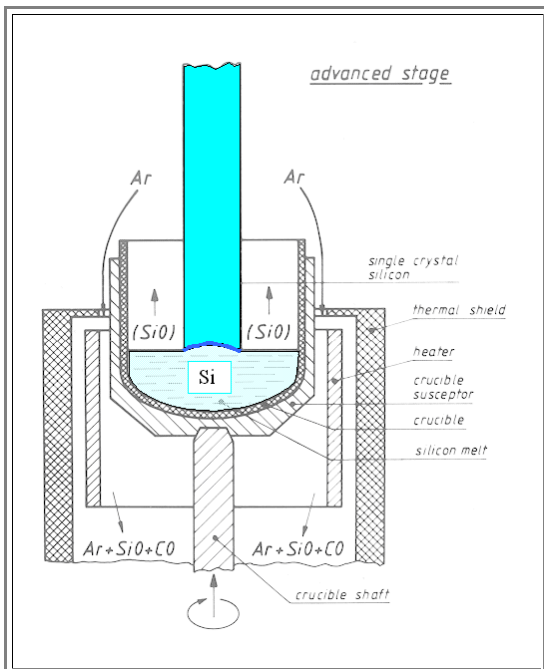
There are a few major points in Crochalski (or **CZ**) crystal growth that shall be outlined briefly. In particular, we will discuss

- The "**Dash**" process (or "necking") for the production of dislocation free crystals.
- **Bulk microdefects**, or the question of what happens to the point defects present in thermal equilibrium upon cooling down?
- **Convection** in the melt and the incorporation of oxygen.



Dash Technique

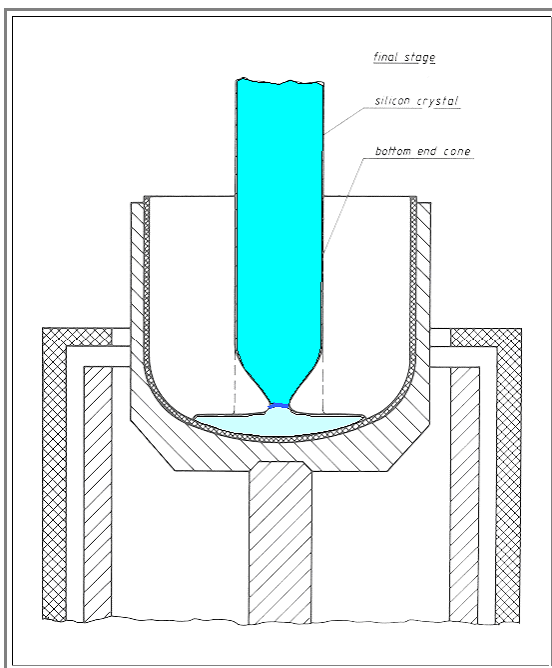
- How do you obtain perfectly dislocation free crystals? Always considering that your seed crystal may not be dislocation free, and even if it is, the unavoidable thermal shock upon dipping the seed crystal in the melt will almost certainly produce some dislocations (this is real easy at temperatures close to the melting point).
- You just do two things:
 1. Reduce the strain as driving force for dislocation movement and multiplication, and
 2. Make sure that existing dislocations are never lined up in growth direction, so that they will sooner or later terminate at the surface.
- Both points are satisfied if you just make the first part of the growing crystal very thin, i.e. you form a "neck" as shown on the left.
- This technique was pioneered by W.C. **Dash** in **1959**. It was one of the many "little" inventions necessary to allow Si technology.



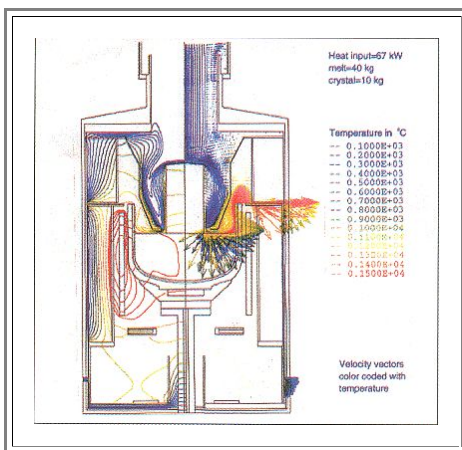
Bulk Micro Defects

- Just after crystallization, the crystal is at a temperature close to the melting point, and thus contains the maximum equilibrium concentration of point defects (vacancies and interstitials). As it cools down, the equilibrium concentration goes down to practically zero at room temperature.
- But how? In normal crystals the point defects disappear at grain boundaries or dislocations - whatever internal sinks there are. In dislocation free single crystalline Si there is only the surface available as sink - and for most point defects, the surface is far away and will never be reached by diffusion.
- What will happen is that some, if not most, of the point defects form small clusters or agglomerates, the so-called bulk micro defects, (**BMD**), **COPs** (crystal originated particles or pits), or, as they were called in the eighties, swirl defects.
- They are unavoidable, and they are bad for **ICs**. All you can do is to try and make them very small. Manipulating these defects is one of the recurrent themes in **Si** crystal growth, made difficult by the fact that - quite ironically - we know far less about point defects in Si than in metals.

Convection Effects



- Even if crystal and crucible would not be rotated, there would be convection in the melt, driven by density gradients coupled to temperature gradients. The only way to totally avoid this would be to switch off gravity. This is the reasoning behind all those " **Very Important**" experiments with crystal growth in space.
- Don't fall for it. First, you also can suppress convection with a strong magnetic field (its actually done for large diameter crystals) at an incredibly tiny fraction of the prize for space-grown crystals, and secondly, why should you suppress convection?
- Well, uncontrolled convection is bad for many reasons. Lets just look at one: The melt streaming by the crucible walls dissolves some **SiO₂** and later deposits it into the crystal. The **O** - concentration then is not only higher compared to convection-free melts, but its radial distribution mirrors the flow pattern - and nothing is worse than inhomogeneities in **Si** crystals.
- Does this mean that you do not want convection? No - all it means that want to have convection that is precisely right for your crystal at every stage of its growth. If you have that, you can control **O** - concentration and keep the doping concentration about constant through the length of the crystal despite the [effects of segregation](#).



Science and Art

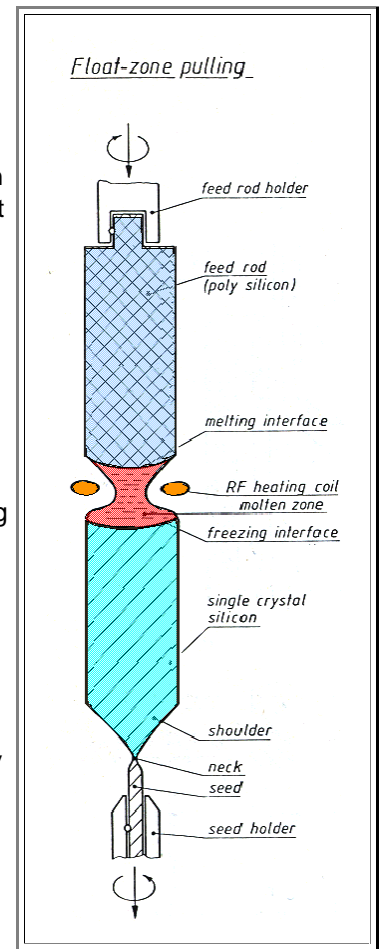
- So, in growing a large **Si** crystal, you have a number of buttons to fiddle with. which you must turn in the right directions as your crystal grows. Consider the major parameters:
 1. Growth speed,
 2. Rotation of the crystal,
 3. Rotation of the crucible - same or opposite direction,
 4. Magnetic field (strength and direction)
 5. Power fed to crucible.
- Changing anyone of this parameters a little bit will influence everything in a way not easily cast into formulas - diameter of the crystal, **BMD** size and distribution, **O**-concentration and distribution, dopant concentration and distribution, and much more.
- Growing good crystals certainly was (and to some extent still is) an therefore an art.
- Only in the nineties of the **20.** century, theory and simulation progressed enough to be of help. Now it is indispensable.
- The picture on the left shows one result: the complete temperature distribution in the whole machine (including, of course, melt and crystal) and the flow field in the melt (little arrows). Now you can change some parameters and see what will happen without actually growing (and later destroying for analysis) an expensive crystal.

Float Zone Crystal Growth

Advanced

The basic idea in float zone (FZ) crystal growth is to move a liquid zone through the material. If properly seeded, a single crystal may result.

- The method was first used for purification (**zone melting**), taking advantage of the small [segregation coefficients](#) of many impurities. The impurities contained in the feed material would then prefer to remain in the melt and thus could be swept to the end of the feed stock.
- If properly done, the newly crystallizing material could be obtained as a single crystal. Again, it was the **Siemens AG** that pioneered the use of FZ crystal growth for the production of high-quality Si single crystals.
- Since the melt never comes into contact with anything but vacuum (or inert gases), there is no incorporation of impurities that the melt picks up by dissolving the crucible material as in the [CZ crystal growth method](#). This is especially true for oxygen, which can not be avoided in CZ crystal growth. FZ crystals therefore are always used when very low oxygen concentrations are important.
- The problem of FZ crystal growth is clear, however, if one looks at the drawing: How do we keep the liquid Si from just collapsing? If it would only be held in place by *surface tension*, the maximum diameter of crystals possible in this way would be about **20 mm** - not very useful. There are, fortunately, other stabilizing mechanisms, and drawing the liquid zone through a "hole" - as indicated - also helps. Still, for large diameter crystals the difficulties grow rapidly and FZ crystal growth is rarely (if at all) used for diameters larger than **150 mm**.



Jan Czochralski



Advanced

While everybody in the semiconductor world knows the "*Czochralski grown crystals*", (almost) nobody knows Jan **Czochralski**.

- He was born **1885** in Kcynia, a small town in Western Poland, which then was part of Prussia; i.e. the German empire. Around **1900** he moved to Berlin, **1907** he worked as an engineer for the "Allgemeine Elektrizitäts Gesellschaft (AEG), a formerly large and famous company, defunct since about **1980** (the last remnants were swallowed by Daimler-Benz). During this time he also studied Chemistry and Metallurgy at the "Charlottenburg Technical University" and Fine arts at the Berlin university.
- **1917** he organized the well-equipped research laboratory of "Metallbank und Metallurgische Gesellschaft"; and he was the director of that institution until 1928.
- **1929** he moved back to Poland to become a Professor of Metallurgy and Metal Research in the Chemistry Department of the Technical University in Warsaw.
- All the time he was involved in metallurgical research, active in many scientific organizations and advising companies and other professional entities. The second world war put an end to all this - J. Czochralski was active in supporting his co-workers and the polish underground army.
- After the war he returned to his native town of Kcynia where he ran a small drug firm. He died in **1953** and is buried in the family tomb in Kcynia.

He was quite well-known for many contributions in metallurgy, especially for his studies of the velocity of metal crystallization, and, of course, for his method of monocrystalline growth, published in **1916**.

- "**Czochralski grown crystals**" are a cornerstone of modern material science, and J. Czochralski will be remembered for his insights leading to this method.
- But like many scientists, he was also interested in literature, music and painting and not afraid of branching out, as witnessed by his activities after the war.
- He is one of the most famous Polish scientists and the Foundation for Materials Research Development in Poland established the Prof. Jan Czochralski Gold Medal for achievements in materials science.

Note added after consulting an Internet article from *Dr Pawel Tomaszewski*. (<http://rekt.pol.lublin.pl/~ptwk/art2.htm>)

- What I did not know was that the discovery of the Czochralski crystal growth methods was one of those **accidental discoveries**; here a quote from the article mentioned above

"DISCOVERY OF THE CZOCHRALSKI METHOD"

The Czochralski method of growing single crystals brought Jan Czochralski his greatest publicity. The method was developed in **1916** and was initially used to measure of crystallization rate of metals. The method was developed as a *result of an accident* and through Czochralski's careful observation.

One evening he left aside a crucible with molten tin and returned to writing notes on the study carried out on a crystallization study. At some moment, lost in thoughts, instead of dipping his pen in the inkpot, he *dipped it in the crucible* and withdrew it quickly. He observed then a thin thread of solidified metal hanging at the tip of the nib. The discovery was made!

The nib slot, in which crystallization was initiated, was replaced by a special narrow capillary and in some cases by a seed of the growing crystal. Czochralski checked later that the crystallized wire was a single crystal. The crystals obtained in that way had diameters of about a millimeter and lengths up to **150 cm**. Czochralski published a paper on the study of the rate of crystallization of tin, zinc and lead, and the maximum rate of pulling of a crystal was recognized as the characteristics of the crystallizing material (Ein neues Verfahren zur Messung des Kristallisationsgeschwindigkeit der Metalle [A new method for the measurement of crystallization rate of metals], Z. phys. Chem. 92, 219-221 (1918); the paper was received in the editorial office on 19 August 1916)."

Box Isolation Technique

Advanced

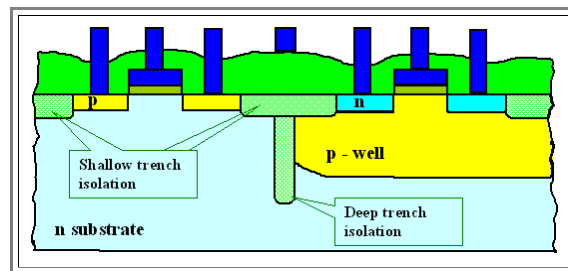
Advanced devices nowadays use a box isolation technique. However, the accepted name for that is "**shallow trench isolation**" or **STI**.

- If the trench is "**deep**" you may still call it **STI**, or possibly DTI, but this abbreviation may also mean "double trench isolation" - just get used to the fact that the semiconductor industry moves so fast to have time for standardizing those things and to force people to stick to it.

How is it done? And why do we need it? Let's just look at this issue very briefly.

First, we note that all the problems with **LOCOS** already mentioned in the [backbone](#) get rapidly worse as dimensions get smaller. On top of that, new problems develop, and the process flow becomes increasingly complicated (and expensive).

- Many problems would disappear if you just would etch a suitable "hole", i.e. a trench in your substrate wherever you need isolation, and then fill it with oxide.
- Or **CMOS** structure [from before](#) then would look like this:



- Of course, nobody would superimpose a shallow and a deep trench isolation as shown (you just would have the deep one and move the transistors closer together), but the picture illustrates the point nicely.

- So why wasn't it done long since?

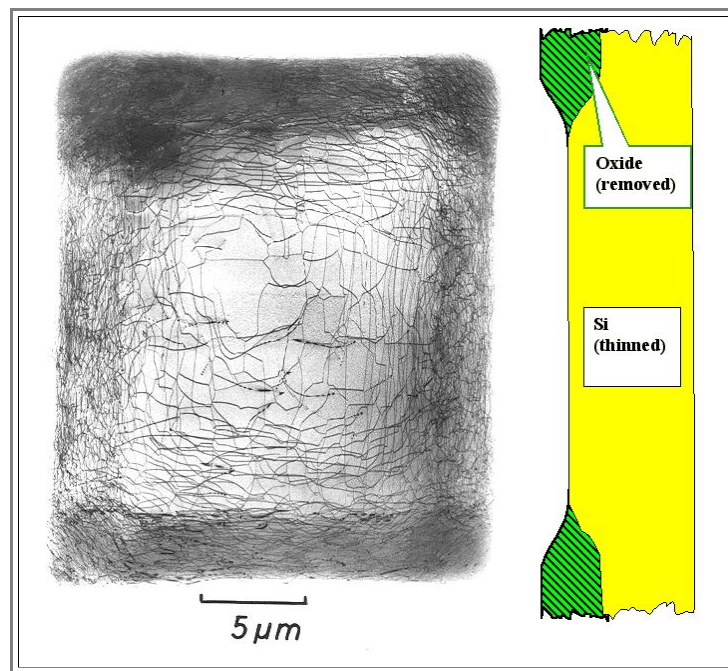
Because it is neither easy to etch the required trenches, to fill them with (high-quality) oxide, and to planarize the surface.

- The latter point is the key: Whatever process you use to **fill** the trenches with oxide, after your oxide deposition you have oxide everywhere, and you must take it off again wherever you don't want it, i.e. outside the trenches
- If you wonder why you **fill** the trenches with oxide, and why you don't use simple **thermal** oxidation anymore, you missed some the essentials! Figure it out yourself; it is enough to consider what would happen if you start thermally oxidizing your **Si** after you etched the deep trench.

TEM of Oxide Edge Dislocations

Advanced

Here is another picture of oxide edge dislocations, this time showing the complete area of the field oxide

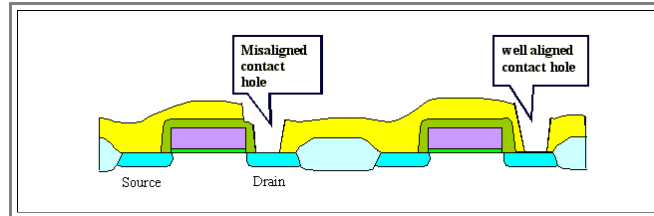


- The graphics on the right indicate the cross section through the specimen (the oxide was removed during preparation). The areas of **Si** formerly covered by oxide are now so thin that they are totally transparent to the electron beam - the dislocated areas seem to float in empty space.
- You may wonder how pictures like this are taken. The link leads to a chapter from the "Defects in Crystals" Hyperscript explaining how [transmission electron microscopy \(TEM\)](#) can be used to image defects in crystals.
- However it works, this picture is special. In any **TEM** image, the electron beam is passed through the specimen, and pictures with good resolution can only be taken if practically all electrons emerge from the specimen backside with the same energy they had before entering the specimen
- We know that electron beams do not travel very far in solids (the electron beam inside your monitor does not only not get out, after all, but deposits its energy into a rather thin layer of luminescent material); if they are supposed not to even slow down, the specimen must be very thin - say **0,5 μm** at the most for a conventional **TEM** with a beam energy of **(100 - 150) kV**.
- This would mean that the specimen must be thinner than the structure it is supposed to contain - only the surface near parts of it would show.
- This problem was overcome around **1976** by using the high-voltage **TEM (HVTEM)** of the Max-Planck-Institute for Metal Physics in Stuttgart - a monster that commanded a beam energy of **650 kV** which was enough to look through samples of **(2 - 4) μm** thickness. The picture shown above (and some of the other ones in the Hyperscript) were taken with this machine and therefore show the full richness of the defect structures.
- Very few **HVTEMs** are in use - they are too expensive (the newest one in Stuttgart costs about **1,5 · 10⁷ DM**), so pictures like the one above are rare.

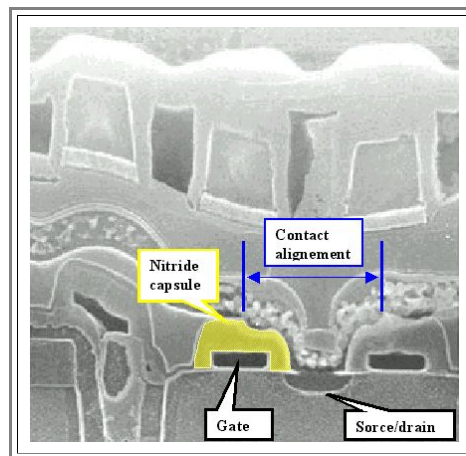
FOBIC Process

Advanced

- Consider making a contact hole through the **CVD** oxide that was deposited over the finished transistors.
- Since the source/drain area are as small as possible, you have to make not only a small hole *but you must align it precisely* relative to the gate stack of the transistors.
- Small misalignments would produce a short circuit between the gate electrode and the **Al** that will be put in the contact hole eventually.
- Now cover the whole gate stack with **Si₃N₄** that will resist whatever etching procedure you use to remove the oxide in the future contact hole. You gain a lot in your "**process window**" for the contact hole - misalignments don't matter as much any more as shown below.



- That looks pretty good - but there is a prize to pay:
- First of all - how do we make the nitride encapsulation? And don't forget; **Si₃N₄** must never come in contact with **Si** - you always need a thin layer of buffer oxide underneath (not shown in the picture). There is quite a bit of added process complexity!
- The topography gets worse. The aspect ratio of the contact hole - the relation between depth and diameter - increases and with it the problems of filling it with **Al**.
- Still, starting around **1987**, **FOBIC** was used and helped to get the next generation onto the market.
- The picture below is a cross section through a **16 Mbit DRAM** memory cell shown before. It shows one of the contacts to a transistor (which is connected on the other side to a trench capacitor). The **FOBIC** structure has been outlined; it is clearly visible.



Spiking and Epitaxial Si in Contact Holes

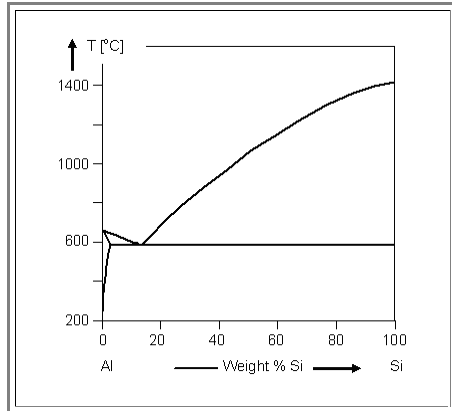
Spiking

Advanced

When we make a contact to the structures in the Si, for example to source and drain regions, we first cover everything with an insulator - **SiO₂** - and then make contact holes in the proper places. Use the [link](#) to refresh your memory.

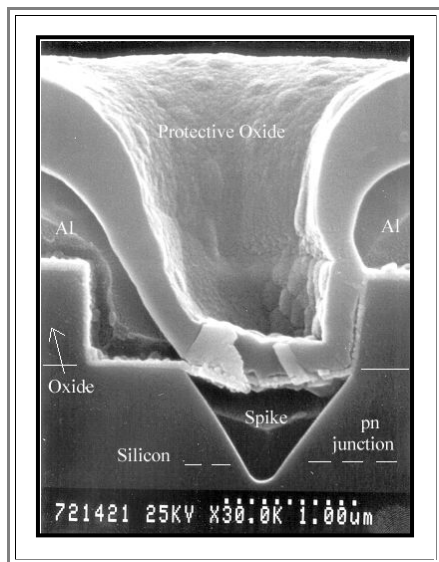
Next we cover everything with a metal - if we stay simple we just use **Al**. We have a lot of other processes after that, and we will have to heat up the wafer to some extent for doing whatever needs to be done.

What happens if we bring **Si** in contact with **Al** (or any other metal, for that matter) and heat it up to some extent?



- Well - consult the phase diagram, and it will tell you what you should expect at whatever temperature you chose. Here is the **Al - Si** phase diagram.
- What it tells us is that somewhat below **600 °C** we will have an eutectic.
- In other words, above the eutectic temperature, "things" will melt. Consult the [chapter about phase diagrams](#) if you don't know exactly what "things" means.
- In yet other words, after **Al** deposition, we must not raise the temperature above the eutectic temperature ever. In order to stay on the safe side, we must even keep it below about **500 °C**.
- This is not so good, but something else is worse: The solubility of **Si** in **Al** around **500 °C** is finite - about **2 %** one would estimate - , while at room temperature it is practically nil.

So what happens if you heat up **Al** in contact to **Si**? The **Al** will try to incorporate some **Si**; it will sort of "suck it up" from the **Si** substrate.



- In more scientific terms we talk about **Si** atoms diffusing into the **Al** (and **Al** atoms diffusing into the **Si**); so temperature dependent quantities like diffusion coefficients are involved.
- If we now have some of the **Si** in the **Al**, it must be missing somewhere else; obviously right below the **Al-Si** interface we must expect some "missing" **Si**.
- If we are lucky, the "missing" **Si** is uniformly distributed, i.e. the whole surface of the **Si** moved down a bit. However, Murphy's law (What can go wrong will go wrong) applies, and on occasion all the **Si** moving into the **Al** comes from one rather localized spot at the interface - we get a "**spike**". This is shown very drastically in the picture.
- At the same time, **Al** will diffuse in the room left by the **Si** - our spike is filled with **Al** and we have a short-circuited **pn-junction**!

This ain't so good. Remember: One spike / short circuit in just one of the > **50 million** or so contact holes will kill the whole **IC**. We must fix that problem.

- Easy, you say. (Do you see the obvious solution?). We do not use pure **Al**, but **Al** already containing some **Si**, so it does not have to "steal" **Si** from the substrate to meet its solubility needs at higher temperatures.
- Right. That is exactly what we will do. We use **Al** alloyed with **0.5% - 1 %** of **Si**. No more spikes will form, and as the process engineer in charge you can sleep well again at night.

Really? Yes, you will. For a few years at least. But then you slumber will become unruly again, because as dimensions shrink, you run into new problem.

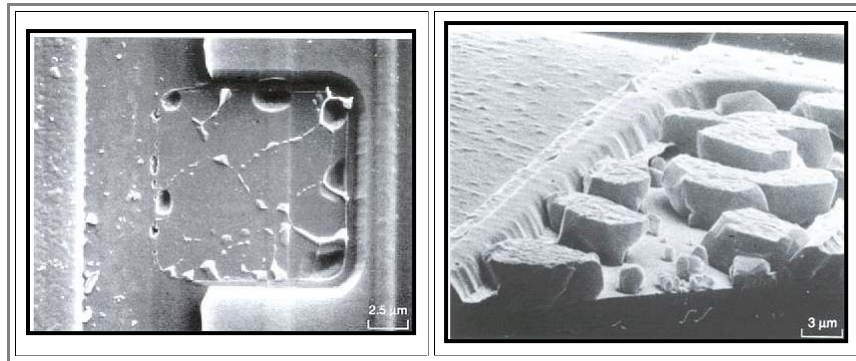
Si Precipitation in the Contact Hole

You use **Al(0.7 % Si)** as your contact material. The dictate of the phase diagram with regard to the Aluminium needs of **Si** at your highest temperature are met; no spiking will occur.

- How about the needs of the **Al** at room temperature? They are not met, because **Si** solubility at room temperature is negligible.
- The phase diagram now dictates that **Si** precipitates should form. This will need some nucleation and may be kinetically difficult in the bulk of the **Al**, but at the **Al-Si** interface we already have **Si** and nucleation is easy.

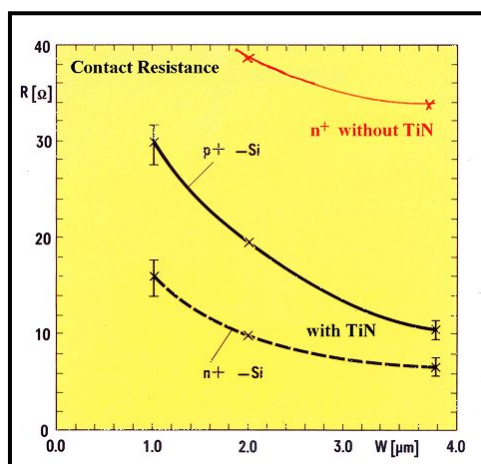
In essence, we must expect the reverse of the "sucking **Si** into **Al**" process to take place. The **Al** will "spit out" its surplus **Si** and deposit it right at the interface

- Si** precipitates will grow *epitaxially* (there is no reason why not) on the **Si** below the **Al**. We must expect to find some **Si** islands on top of the **Si**



That is exactly what we see in the pictures above.

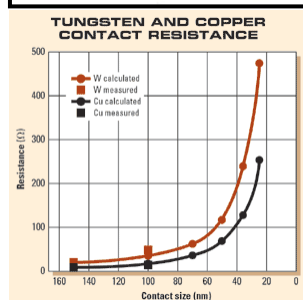
- In the left hand picture we have a nice balance of a few shallow spikes and some **Si** precipitates following the contours of the grain boundaries in the **Al** (which has been etched off). This is not surprising, because precipitation is a diffusion process, after all, and diffusion along grain boundaries is faster than in the bulk.
 - The right hand picture shows rather large precipitates, almost taken up all the space there is in the contact hole.
- The question, of course, is: Does it matter? We still have some **Al** in contact to our **Si** substrate, and then we have **Al** in contact to the **Si** precipitates and the **Si** precipitates in contact to the **Si** substrate. Plenty of possibilities to run a current through the contact.
- Right, but **Al** in **Si** is an acceptor, and we must expect the **Si** precipitates to be saturated with **Al** and thus heavily **p**-doped.



- No big problem as long as the **Si** below the contact is **p**-doped, too, but a big problem for **n**-doped **Si** and small contacts.

In this case part of the contact area is now a **pn**-junction, blocking current flow in one direction.

It is clear that the **contact resistance** always increases with decreasing contact area, but if we have **p**-doped **Si** precipitates, we have only part of the geometrical contact area for the contact to **n**-type **Si**.



What we get as a function of the contact hole size for the **n**-doped case is shown in the figure in red.

The message is clear: Contact resistance will be too large at some point. So what do you do?

Without the **Si**, you get spikes, with the **Si** you get precipitates and a contact resistance that is too large.

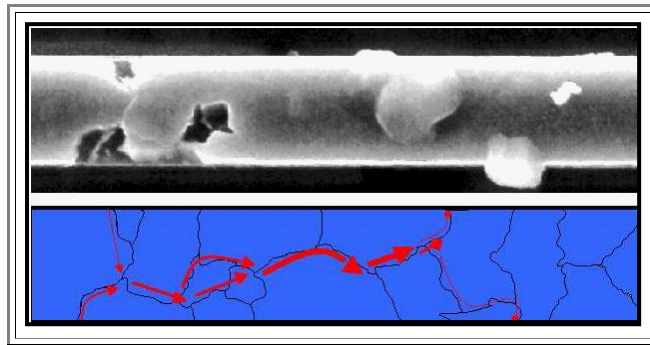
- There is no choice anymore: You need a new material; in this case a diffusion barrier between the **Si** and the **Al**.

- Introduce, e.g., a thin layer of **Ti/TiN** between the **Si** and the **Al**, and your contact resistance problems are solved, as seen in the figure. If you also throw in [contact hole filling with W-CVD](#), you now have a structure like the one on the [title page](#).
- Except, of course, that now you have to worry about how to make and structure this layer. How to measure how thick it is, and if it has the required properties (assuming that you know that).
 - You must worry if the new diffusion barrier interferes with the reliability of the metallization (after all, [electromigration](#) in the metal is one of the major causes of premature device failures), and if can keep the additional processes cheap.
- In other words. Introducing a new material is a long and cumbersome process.
- And a frustrating one. Nobody tells you, that **TiN** is it! You (meaning you and your team) find that out yourself. And in that process you will test many materials, most of which will not be the right ones.
 - Out they go - and with them many hours of your time and a lot of (other peoples) money.

Electromigration

Advanced

- ▶ The picture below shows an **Al** circuit line of an **IC**; it is about **1 μm** across.
 - It has been "**stressed**" for some time by running a **DC** current through it at temperatures and current densities somewhat higher than what it would encounter in normal operation. The minus pole was on the left. The current density may have been in the **10^5 A/cm^2** region.
 - This is a hell of a current density - several orders of magnitude larger than what you would feel comfortable with in normal wires.
- ▶ The conductor line obviously has suffered from this treatment. In its left part - close to the minus pole - holes or pits have formed, and in its right part - close to the plus pole - we have extrusions or hillocks.
 - Obviously, substantial amounts of **Al** have migrated to the right: leaving back holes and generation hillocks. and
 - Considering that a huge number of electrons was driven from left to right by the current source, it almost looks as if those electrons have pushed some **Al**-atoms along.
- ▶ This is indeed what has happened. The phenomena of **electromigration** is, in general terms, a drift in the diffusion currents induced by the "**electron wind**"
 - Remember that there is always some self-diffusion, but without any driving forces, the net diffusion current in one direction is always cancelled by the exact opposite current in the other direction.
 - Exactly how those electrons push atoms about, is still a bit obscure, but the effect is clearly there, and appreciable in finite time spans if the current densities are extremely large.



- ▶ You thus would expect that electromigration scales with the [self-diffusion coefficient](#), and thus is more severe in materials with a low melting point.
 - Yes, but classical self-diffusion is a bulk property. In our poly-crystalline **Al**, we also have diffusion along grain boundaries, which is generally much faster than bulk diffusion.
 - A more detailed analysis of the picture above (which, incidentally, are from the thesis work of Dipl.-Ing. Wedemeyer in **1998**) that includes the grain boundaries clearly demonstrates that grain boundaries indeed provide the main pathway for the movement of the **Al** atoms.
 - ▶ At the holes, the cross-section of the **Al** wire is reduced, i.e. its local resistance **R** is larger, and the locally dissipated power **$L = I^2 \cdot R$** at a constant current **I** goes up.
 - The wire heats up at the holes, electromigration becomes more severe, holes get bigger, and soon you have a complete interruption of the current path and thus the end of your **IC**.
 - ▶ We have a problem. As dimensions shrink by a factor of κ , the cross-sectional area of the **Al** lines tends to shrink by κ^2 , but the current only goes down linearly with κ .
 - With shrinking dimensions current densities thus tend to go up, and electromigration in pure **Al** (or **Al** with some **Si** to prevent [spiking](#)) would have made integrated circuits with a life time of **10** years simply impossible in the **80**ies of the last century.
 - You must do something. Avoiding grain boundaries roughly parallel to the wire helps, but how are going to do this? Well, there are some tricks which help to create the "bamboo structure" you want, but this would not be sufficient.
- So you do what you **always** do in this kind of situation: You add some dirt to your **Al** and hope for the best.
- ▶ Maybe, the right element "poisons" the grain boundaries, making grain boundary diffusion for **Al** atoms more difficult. Or maybe.....
 - Having a "theory" or just a hunch to work from helps; but not too much. Doing is what really helps. And yes, if you add about **0.5 %** of **Cu**, electromigration becomes less severe. This is also true for other elements, but **Cu** is what is mostly used.
 - However, if only a tiny fraction of that **Cu** ever diffuse into the **Si**, it will kill your device. Also, your etching process for **Al** may suddenly not work anymore. And your **Al** metallization now might be more sensitive to corrosion. And the adhesion of the next layer is now different . And....

- You get the drift: There will always be a lot of demanding work out there for highly qualified people like you!

Doping Trench Walls

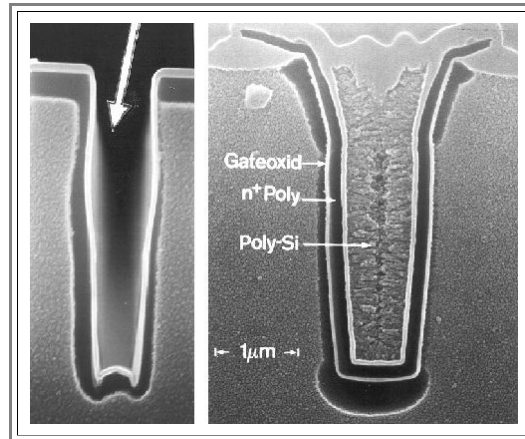
Advanced

You have a trench (i.e. in [reality a hole](#)) with a diameter of about **1 μm** and a depth of **5 μm** . The **Si** around it serves as one of the capacitor electrodes and must be doped with something to a depth of about **0,3 μm** so it is highly conductive.

This task came up as an unprecedented challenge to process engineers in the context of the **4 Mbit DRAM** generation around **1985**. How do you do it?

First you try what you have - ion implantation - and hope for the best. Since the ion beam does not impinge at right angles on the specimen, but somewhat off (say about **10°**) for various reasons, there is hope.

Indeed, it works - see below:



With a special etchant the doped region is etched off, leaving a dark rim around the trench (which was not very perfectly etched at this point in time).

The arrow in the left-hand picture marks the major direction of the ion beam, and the left-hand side of the trench is nicely doped, as one would expect. The right hand picture shows a fully processed trench capacitor for measurements of the electrical characteristics.

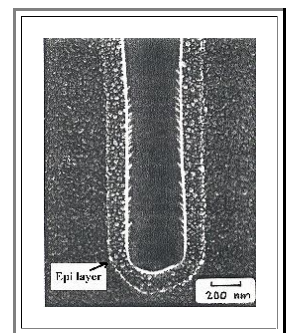
But some implantation also took place on the unfavorable side - there is a certain spread in the beam angle, some ions are reflected or scattered off their original direction - whatever happens, some ions hit the wall. If you turn the wafer during implantation, it might be possible after all, to obtain decent homogeneous doping.

But in the end, ion implantation was not used. Too expensive, too dependent on the precise trench geometry (which changes all the time), simply not good enough for prolonged mass production.

OK, so you try [epitaxy](#). Simply deposit the doped layer into the trench as you want to have it.

Allright - it works. Some problems are encountered around the opening of the trench, but nothing that couldn't be solved eventually. The picture shows part of a trench with the doped epi layer; this looks good.

But we don't use it for production! The reason is simple: Too [expensive](#), too hot - and we can do it better in a different way.



How? Well, let's deposit a "[spin-on glass](#)" (**SOG**) containing the dopant, hope that it will completely fill the trenches, and then diffuse the dopant out of the **SOG** into the **Si**.

Spin-on glasses are used a lot, but do they fill up fine holes? The only way to find out is trying it.

To cut a long story short - it worked. But not reliably enough; didn't make it to production.

What finally made it to production is a somewhat similar approach, except that the **SOG** was replaced by an **As** doped [TEOS-oxide](#).

Not an extremely nice process (the coating of the **CVD** tube is **As** contaminated and highly toxic), but without many unknowns and reproducibly.

Is there something even better?

- Yes, there is a simple solution to the trench doping problem: *Don't do it*. Oxidize the trench walls and deposit a layer of highly doped poly-**Si** as the capacitor electrode. Simple. Elegant.
- Except that now you have to produce the very thin capacitor dielectric on poly-**Si**. [Thermal oxidation](#) of poly-**Si** will not be good enough - now you need [ONO](#). The result was already shown in an [illustration module](#)

Dry Etching - Some Special Issues

Basic Dry Etching Modes

Advanced

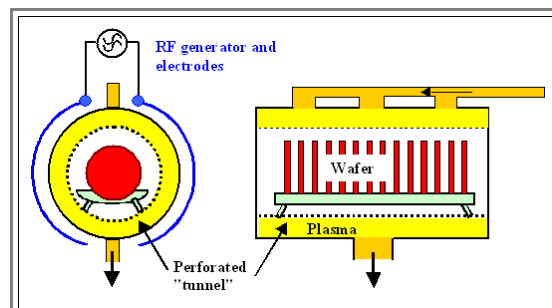
There are four basic modes of dry etching

1. **Etching in a gas** (**Chemical dry etching, CDE**). Rarely used, but possible. As an example, with ($\text{HF} + \text{N}_2 + \text{H}_2\text{O}$) vapor, some **CVD** oxides can be etched about **10** times faster than thermal oxide - a selectivity not achievable with wet etching. Like all chemical etches it is isotropic.
2. **Chemical etching in a plasma**. Whatever is produced in the plasma then must **diffuse** to the wafer, i.e. the physical component of large kinetic energies is absent. Like all chemical etches it is isotropic. This is a process widely used for hard-to-wet-etch layers, in particular photo resist (which is simply "burned off" in an **O**-plasma). This was the first dry etch process to hit production.
3. **Ion beam etching** (called **RIBE** for "**reactive ion beam etching**"). The etching process with high energy ions extracted from some plasma source is purely physical/mechanical. **RIBE** has all kinds of problems and is rarely used in its undiluted form.
4. **Chemical-physical etching** called **RIE** for "**reactive ion etching**". Here we mix everything from above and this is where it becomes complicated but powerful. This is what we use for all critical etch processes

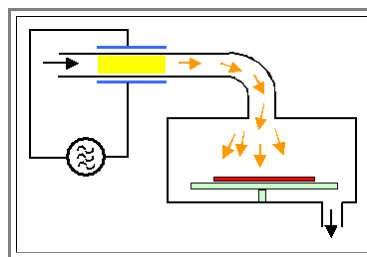
Basic Reactor Types

Here are a few basic reactor types. You always go for batch processing, i.e. for etching several wafers (ideally **25 - 100**) at once, if it is possible (which ever more often it is not). We skip the reactors for **CDE** and **RIBE** because these etch modes are not so important.

- Chemical etching in a plasma**. There are two basic types of reactors: The **barrel reactor** and the "**down stream**" reactor. In both types the plasma does not extend to the wafer and the etching species must diffuse to the wafer.



Basic set-up of a barrel reactor - the work horse for resist removal.

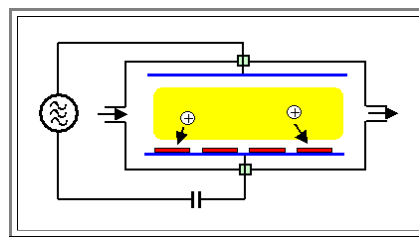


Basic set up of down stream reactor.

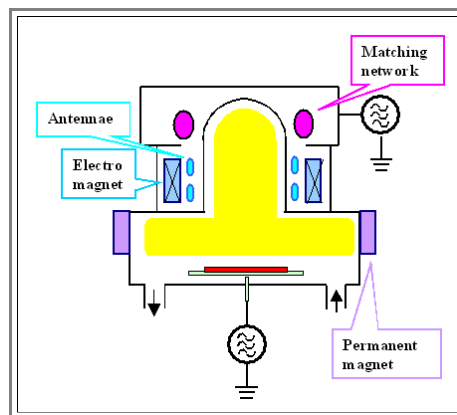
The plasma in both cases might be produced by high frequency excitation, or even with microwave induced discharge. The pressure is relatively high, around **100 Pa**.

- Besides photo resist etching in an **O**-plasma, **Si**, **SiO₂** and **Si₃N₄** might also be etched (isotropically) if some **F**-bearing gas is used.

For reactive ion etching, some kind of **parallel plate reactor** or some kind of plasma-source reactor is used. Designs might look like this:



A basic parallel plate reactor - quite similar to the [basic set-up for sputter deposition](#).

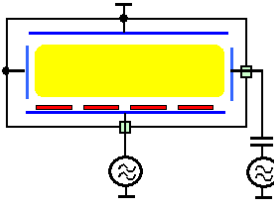
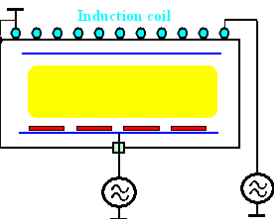


A rather complicated plasma reactor, just to give an idea of where plasma etching is going

Basic Etching in One Reactor Type

Let's just look at etching in parallel plate reactors for some relevant examples.

Mode	Basic configuration	Advantages	Disadvantages
RIE Reactive ion etching		<ul style="list-style-type: none"> * Large anisotropy * Good structure transfer 	<ul style="list-style-type: none"> * Low selectivity * Low etch rate * Surface damage
Anodically coupled plasma etching		<ul style="list-style-type: none"> * High selectivity * High etch rate * Low surface damage 	<ul style="list-style-type: none"> * Tendency to underetch mask (bad anisotropy)
MRIE Magnetically enhanced RIE		<ul style="list-style-type: none"> * Large anisotropy * High etch rate * Reduced surface damage * Etching relatively independent of loading 	<ul style="list-style-type: none"> * Homogeneity is a problem

TRIE Triode reactive ion etching		* Large anisotropy * Increases process optimization potential with generators	* Increased costs and complexity
TCP Transmission coupled plasma etching or Inductively coupled plasma etching		* Etching relatively independent of loading * Very large etch rates because of high plasma density	* Constructive problems (no ferromagnetic steels can be used)

Well, you get the idea. And if you do not understand exactly how this works (consider the variations in what is grounded and if there is a capacitor), don't worry: It *is* quite involved. Let's just look at the first two examples:

Here is **RIE**, the first example given in the table.

- In the case of pure **RIE**, the high frequency voltage is *coupled capacitively* to the lower electrode that carries the wafers. The upper electrode is grounded, and together with the chamber it has a much bigger surface than the lower electrode.
- This will lead to a *negative charging* of the lower electrode (the reasons for that are tricky and due to the fact that no **DC** currents can flow, and that the mobility of the negative electrons and positively charged ions is very different).
- The positive ions thus are strongly accelerated before they hit the substrates to be etched; typically they have energies > 100 eV.
- If we keep the pressure low enough, there is hardly any scattering of the ions, and they impinge nearly vertically on the substrates.

Now let's look at the second case, *anodically coupled plasma etching*.

- It looks rather similar to the first case, we just changed the connection of the **HF** generator and the ground.
- This simply reduces the negative charge on the lower electrode, and thus the energy of the ions hitting the substrates.
- In addition, we will keep the pressure rather high, making sure that the ions accelerated in the direction of the substrate have many collisions with gas molecules and thus bombard the substrate from all directions. This also reduces their energy, and etching is less "physical" and more "chemical". We gain in selectivity and loose anisotropy.

But now we will stop. While you should marvel at the complexity of plasma etching, you should not despair. It was people like you - and not the Noble prize winners in physics - who made it work and continue to improve the technique.

CaF₂ Lenses

Here is an article from the Semiconductor International from February 2002

Advanced

At a Glance

- Calcium fluoride will likely remain the lens material of choice for **157 nm** lithography tool designs, engineers having come up with various tricks to offset birefringence problems. But optics and system developers have their work cut out for them.

157 nm Optics Demand a Bag of Tricks

Aaron Hand, Managing Editor -- *Semiconductor International*, 2/1/02

Calcium fluoride. That material has evoked a roller coaster of hopes and fears in a chipmaking industry gearing up for 157 nm lithography and the finer features that it enables. Lithographers have for years relied on readily available fused silica for system lenses, but the material does not transmit well at 157 nm. Enter CaF₂, identified as a possible savior to 157 nm technology.

However, the difficulties that come with crystalline materials caused concern that the industry would not be able to get the quantity or quality it needed. But engineers persevered, producing generations of CaF₂ and overcoming various technical barriers, such as distortion and stress-induced birefringence. Various organizations were working together to help CaF₂ suppliers improve their growth processes. Things were looking up. (For a discussion of preliminary life-test results of CaF₂ [Fig. 1], see "Testing Optical Damage for 157 nm Lithography".)

But in May 2001, at the 2nd International Symposium on 157 nm Lithography, a bomb was dropped. John Burnett, a scientist at the National Institute of Standards and Technology (NIST, Gaithersburg, Md.), reported that CaF₂ suffers from an unexpected intrinsic birefringence at 157 nm. Its birefringence has been measured at ~11 nm/cm, much larger than the ~1 nm/cm birefringence that 157 nm lithography can handle. Such birefringence would cause significant wavefront aberration in the projection optics, putting "showstopper" on the tongues of key players.

By July, at International SEMATECH's workshop on CaF₂ birefringence, the situation was apparently under control. There was relief as researchers showed compensation techniques that could be used to minimize the impact of birefringence. CaF₂ birefringence wasn't going to be a big problem after all.

Today, technologists are surging ahead with optical techniques, aiming to have 157 nm research tools available by 2003 and production tools by 2005. At the same time, though, they are perhaps more realistic about the formidable challenges that lie ahead. Yes, they have come up with a bag of tricks to compensate for the intrinsic birefringence of CaF₂, but those tricks are going to prove difficult and costly, both in terms of time and money.

Although it took the industry by surprise initially, lithographers are relieved that birefringence will not be the showstopper once thought, said Stephane Dana, product manager at ASML (Veldhoven, Netherlands), in charge of 157 nm lithography. Since then, ASML has been taking measurements, along with lens supplier Carl Zeiss (Oberkochen, Germany), and they have become more confident. "We will be able to compensate for it in the optical design," Dana said. "We know we can solve it. The problem is that it puts even more pressure on CaF₂."

Countermeasures will require more effort, time and money, noted Phil Ware, senior fellow, lithography, at Canon U.S.A. Inc. (Irving, Texas). "However, the impact on Canon's 157 nm tool schedule is relatively minor. We expect about a half-year delay for the beta tool (from the first half to the second half of 2003), mainly due to the additional time it will take to grow CaF₂ crystals with different orientations."

Refraining from refractive

Lithography system manufacturers plan to put at least part of their faith in catadioptric designs, combining lenses and mirrors in an optical system to minimize chromatic aberrations. Whereas previous optical systems have been primarily refractive designs, replacing some lenses with mirrors helps to minimize the use of CaF₂. All toolmakers will have to rely on various forms of catadioptric systems, Dana noted. Alternatives — such as line-narrowed lasers and practical alternative crystal materials — have not progressed enough to allow continued use of refractive designs, he said. However, Dana declined to give details of ASML's catadioptric design (for the most part, toolmakers are holding that information close to their chests).

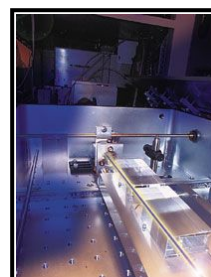
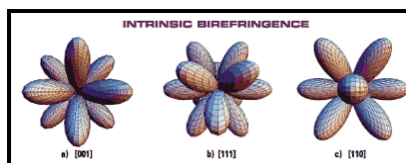


Fig. 1 Testing the life time of CaF₂

Some proposed approaches have featured multiple axes with beam splitters or mirrors, Ware noted. "Canon's approach uses a concentric off-axis catadioptric design with only two mirrors," he said. "This design not only uses fewer elements, but it lends itself to the same kinds of assembly and tuning techniques that we have optimized for our all-refractive optics."

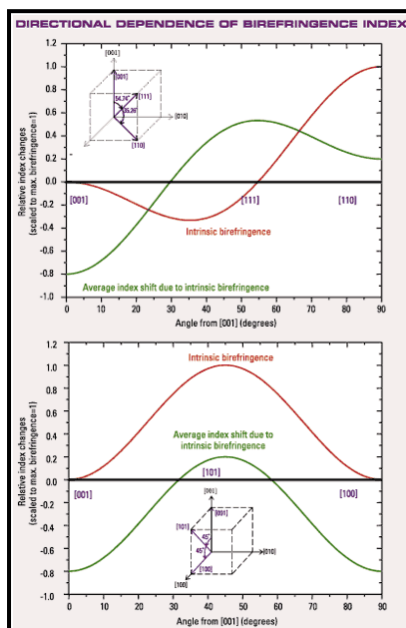
Besides minimizing the use of CaF_2 , optical designers must now consider the orientation of the CaF_2 crystals in their designs. Lithographers have been using lenses made with crystal orientations of [111]. Although clocking these crystals around the optical axis has been suggested as a way to mitigate the birefringence problem, it's not enough. Designs will have to combine [111] crystals with CaF_2 crystals of other orientations to solve the problem more effectively.



2. These 3-D models show the directional dependencies of CaF_2 's intrinsic birefringence. (Source: NIST)

"At this point the designers have stated in the conferences that they will pursue the strategy of catadioptric designs using [001], [111] and possibly [110] clocked pairs, and combining them so that the residuals partially cancel," Burnett said.

The orientation of the lenses is an important factor because intrinsic birefringence depends on the direction in which light propagates through the CaF_2 , which is a function of the cubic crystal direction (Fig. 2). For [111] lenses, birefringence is zero as long as the light doesn't stray from the central direction of the lens. Unfortunately, the angle dependence of this model (Fig. 3) doesn't bode well for lenses with high numerical aperture (NA), which are essential for achieving the resolution necessary for smaller feature sizes. Lenses made with [001] crystals, which are less affected by rays outside the central angle, will particularly help to minimize birefringence, Burnett said.



3. Quantitative values (scaled to 1 at the maximum) demonstrate the directional dependence of intrinsic birefringence as well as the resulting index anisotropy. (Source: NIST)

Canon's design incorporates a combination of [110], [100] and [111] crystal orientations to reduce aberrations, Ware said.

ASML plans to clock elements of [111] and [100] oriented CaF_2 , Dana said. Nikon's primary approach will also be a catadioptric design, combining crystal orientations and clocking strategies, noted Chris Sparkes, director of technology at Nikon Precision (Belmont, Calif.).

Another factor that must be considered, Burnett said, is that the angle dependence of the effect results in an anisotropy of the average index on the scale of the birefringence effect. "This effect also results in wavefront errors and must be considered in the design for minimizing the distortion." The effect has a different angle dependence and different impact on the optical performance, he noted, "so it must be treated separately using the various optical compensation methods."

Since presenting its findings to the industry, NIST has confirmed by measurement the full angle dependence of intrinsic birefringence predicted by the original model, Burnett said, and the model is generally accepted as correct. Optical Research Associates (ORA, Pasadena, Calif.) has incorporated the model into its widely used CODE V lens design software. CODE V can now account for the effects of intrinsic birefringence approximately.

To add to the difficulties, clocking the lenses to compensate for birefringence limits the freedom to clock the lenses to compensate for surface figure and index aberrations, Burnett noted. "This aberration clocking has always been required to build projection lens systems for longer wavelengths," he said. "Restricting this aberration clocking freedom (e.g., to 120°

for [111] lenses) because of correcting for intrinsic birefringence may require tighter surface figure and index specifications."

Quality and quantity

Putting all the additional requirements together only makes matters worse for the supply and availability of high-quality CaF_2 . The quantities available have been an ongoing concern because most CaF_2 suppliers are still on a fairly steep learning curve. CaF_2 is used in conjunction with high-purity fluoride crystals for 193 nm lithography, and has already faced a shortage of materials, noted Gitimoy Kar, director of the Fluoride Crystals Business at Corning Inc. (Corning, N.Y.). Growing sufficiently large crystals is a difficult process, he said.

Now materials suppliers must not only meet increasingly stringent specifications for [111] CaF_2 , but they must grow high-quality crystals in other orientations as well. "Crystal growth is a very slow process; it takes months to grow one ingot," Kar said. "Then you have to multiply the cycle time by the number of orientations you have to grow it in."

Corning acquired the capability to produce CaF_2 in 1999, and has since quadrupled its manufacturing capacity. Other CaF_2 suppliers have worked to expand capacity as well. For example, Schott Lithotec AG (Mainz, Germany) announced last year a third production plant for CaF_2 monocrystals. A subsidiary of Schott Glas, Schott Lithotec was established in 1998, launching its second crystal production facility in 1999.

In an effort to maintain some control over CaF_2 supply, toolmakers are also producing some of their own. Last summer, Nikon Precision noted ramp-up of its own CaF_2 production. Although it would not constitute the majority of its supply, it would be significant in helping the company get the quality it needed. Canon Inc.'s wholly owned subsidiary Optron Inc. (Yuuki City, Japan) inaugurated its CaF_2 crystal manufacturing facility in August, with Canon then claiming itself as the world's largest CaF_2 manufacturer. "When it reaches full capacity over the next year or so, it should meet all of Canon's CaF_2 needs and then some," Ware said. For now, Canon continues to buy crystal materials from multiple outside sources, he noted.

With all the capacity expansions, it seems as if the industry ought to have CaF_2 coming out its ears. Unfortunately, all the capacity in the world doesn't resolve the fact that yields are reportedly in the single digits. Suppliers keep specific yield numbers to themselves — and there is certainly some discrepancy as to what specifications are considered in any yield quote — but yield is "very low," Corning's Kar conceded.

It doesn't help that toolmakers keep tightening specifications and adding measurement requirements, he said. "The industry needs more standardization for crystal specifications," he said. "We spend an enormous amount of time just characterizing materials." But there is certainly room for improvement and progress is being made. ASML's Dana said that the industry needs to see a yield increase of about 25% to get the quantities needed.

"In absence of high yield, people are putting in more furnaces, which is very capital-intensive," Kar said. Although the investment may be worthwhile, it is difficult to continue pouring money into materials for 157 nm lithography when 193 nm lithography has not been generating sufficient revenues. "The industry is in a recession," Kar said. "With all the effort we had put into 193 nm applications, we still need to generate revenues. At the same time, we're being asked to accelerate RD for 157."

Although crystal growers are primarily relying on the standard Bridgman-Stockbarger method, startup Single Crystal Technologies (SCT, Gilbert, Ariz.) has a crystal growth technique that could increase yields to some 90%, according to Kiril A. Pandelisev, CEO and chief science officer (see [Semiconductor International](#), December 2001). Pandelisev contends that improving purification and yield levels is the key to dealing with birefringence issues (see ["CaF₂ Birefringence, Purity and 157 nm Technology"](#)).

Mixing it in

Another option in the battle against intrinsic birefringence is the possibility of combining CaF_2 with other crystal materials. Having shown that CaF_2 has birefringence values that are opposite those of BaF_2 and SrF_2 , NIST researchers believe alternate materials may prove useful in compensating for the effect.

"Our main effort now in this area is exploring the possibility of creating mixed crystals," Burnett said, noting examples of $\text{Ca}_x\text{Ba}_{(1-x)}\text{F}_2$ and $\text{Ca}_x\text{Sr}_{(1-x)}\text{F}_2$. "With appropriate values of x (Ca content) determined from our measured values, these should have no intrinsic birefringence at 157 nm, if high-quality materials can be made."

Several CaF_2 suppliers are working with NIST on this project, according to Burnett. "This approach is appealing since it could eliminate the problem completely," he said. "Such mixed crystals have been successfully grown before for certain ranges of component ratios. However, it remains to be seen whether the other properties such as absorption, index homogeneity, and stress-induced birefringence can be maintained within specifications. Also, given the aggressive 157 nm schedule, the industry cannot count on it, at least for first-generation tools."

Toolmakers seem to be shying away from this approach for compensation, at least for the time being, because alternate crystal materials do not offer the necessary properties. "We really don't see the quality that we need, and certainly not in

the quantity we need," Dana said. "So ASML is not considering this as an option."

Canon and Nikon do not currently plan to use either BaF₂ or SrF₂ for 157 nm tools either. "CaF₂ has introduced its own set of challenges, even after considerable experience in lithography tools," Nikon's Sparkes said. "The use of BaF₂ is even less mature and presents even greater challenges."

The materials are not practical for lens fabrication, Canon's Ware agreed. "Barium fluoride still does not meet the transmission requirements at 157 nm," he said. "It also has 100x greater water solubility than calcium fluoride. This phenomenon would require development of new surface treatment technologies (for instance, we have to use water for lens polishing). ... Based on the information presented to date, we doubt that anyone will successfully use these kinds of materials for full-field 157 nm production lenses."

Whatever the solutions — or, more precisely, whatever the final combination of solutions — may be, it is clear that researchers have their work cut out for them. Intrinsic birefringence certainly complicates system designs and fabrication. Although any hopes of keeping 157 nm lithography on schedule are perhaps a bit optimistic, the semiconductor industry is no stranger to the challenge.

For more information...

When you contact any of the following manufacturers directly, please let them know you read about them in Semiconductor International.

ASML	Canon	Carl Zeiss
Corning	Cymer	Nikon Precision
NIST	Schott Lithotec	Single Crystal Technologies

CaF₂ Birefringence, Purity and 157 nm Technology

CaF₂ has been chosen as the best material to enable the next generation of semiconductor devices. Extensive theoretical and experimental research has now addressed the optical properties and crystal growth characteristics of CaF₂ for 157 nm applications. Intrinsic birefringence, inherited through the crystal structure, and process-induced birefringence are two key topics of this research.

John Burnett et al of NIST have analyzed intrinsic birefringence of CaF₂.¹ Their study indicates that birefringence at wavelengths of 193 and 157 nm can result in variations of 3.6 nm/cm and 6.5 nm/cm, respectively. The calculated acceptable target birefringence value for 157 nm photolithography is 1 nm/cm or less.

To eliminate the effects of this intrinsic birefringence, the scientists have reported that using pairs of lenses cut along the [111] direction with transverse axes rotated by 60° can redirect light to the correct path. One might say that this finding has turned the intrinsic birefringence of CaF₂ into a non-issue.

But increasing the lens stack size intensifies the need for very high-purity, high-quality CaF₂ lenses. Stress-induced (i.e., non-intrinsic) birefringence must absolutely be reduced to as close to zero as possible to prevent unpredictable effects that could render a lens unusable. The increased number of lenses through which the light must pass increases the need for more "perfect" crystals, since the negative effects of lens defects are additive for each new lens used to compensate for birefringence.

Purity thus becomes a primary concern. Present purification processes can generally achieve consistent purity levels of 99.9-99.999%. However, 157 nm lithography using CaF₂ lenses will require crystal purity of at least 99.99999%. With current methods, such stringent purity requirements cannot be met under production conditions.

To resolve the above issues of purity, low to no induced birefringence, and to increase the overall yield of the process, the author has proposed new methods for purification and crystal growth² of CaF₂.

By introducing reactive substances to the crystal starting material while it is in powder or molten state, one can obtain CaF₂ that is several orders of magnitude more pure than that currently available. Growing CaF₂ in plate form also enables one to obtain lens material that is almost birefringence-free. Furthermore, the crystal growth and cool-down cycle take a fraction of the time it would take to grow a conventional cylindrical boule of the diameter needed for the same size lens.³⁻⁴ Better understanding of birefringence, both intrinsic and induced, higher-purity starting material and adopting the plate growth technique promise to pave the road to full realization of the promise of 157 nm technology.

1. John H. Burnett, Z.H. Levine and E.L. Shirley, "Intrinsic Birefringence in 157 nm Materials," 2nd International Symposium on 157 nm Lithography, May 14-17, 2001, Dana Point, Calif.
2. K.A. Pandelisev, "Continuous Crystal Plate Growth Process and Apparatus," Patent No. 5,993,540, Patent No. 6,071,339 and Patent No. 6,153,011; and pending patents on crystal growth and materials purification-related subjects.
3. K.A. Pandelisev, "Purification and Crystal Growth of CaF_2 Having Superior Properties," 2nd International Symposium on 157 nm Lithography, May 14-17, 2001, Dana Point, Calif.
4. K.A. Pandelisev and S. Motakef, "Method and Apparatus for Purification and Crystal Plate Growth of Calcium Fluoride for 157 nm Applications," 157 nm SEMATECH Technical Data Review, Dec. 11-13, 2001, Orlando, Fla.

Some Issues from Advanced Lithography

General

Advanced

- A somewhat better equation than the [one in the backbone](#) for the resolution limit or minimal feature size d_{\min} of an optical system is

$$d_{\min} \approx \frac{k \cdot \lambda}{NA}$$

- λ is the wave length. The parameter k lumps together the effects of, e.g., photoresist response, or reticle properties.
 - The numerical aperture NA can be defined as $NA = n \cdot \sin\Phi$ with n = refractive index of the medium above the photo resist, and Φ = largest angle of converging rays hitting the resist at a "point".
- If we want the ultimate in resolution, we have to work at all three parameters

Wavelength

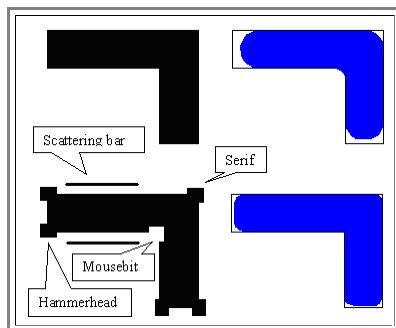
- The visible range of wavelengths extends from about **780 nm** (red) to **380 nm** (violet). Obviously we need to go to even smaller wavelengths in the ultraviolet part of the spectrum if we want to make structures in the **100 nm** region. Obvious, so where is the problem? Well, there are two major problems with this approach.
- *First*, we need a *powerful* and fairly *monochromatic* illumination source, and *second* we need materials to make an extremely good lens from.
- Let's look at the illumination source issue first:
- A *powerful* light source we need because we cannot afford to wait forever before an exposure is finished. The maximum exposure time should be below a second or so, and you simply need intense light for that.
 - *Monochromatic* light we need, because we cannot possibly build a supreme lens for many wavelengths (there are things like chromatic aberration and so on). Taking a small part of the spectrum out of some blackbody radiation (the spectrum emitted by something hot like a light bulb), however, leaves very little intensity.
- The solution lies in going for an intense line in the emission spectrum of some element - mercury (**Hg**) in this case.
- In the **80ties**, the so-called **G-line** at **436 nm** was used (coming from a high-pressure **Hg** discharge lamp). Next came the **I-line** at **365 nm**, and then a **250 nm** line.
 - But that was already pushing the **Hg** lamp to its limits, and it was soon replaced by so-called **DUV** (for deep ultraviolet) **excimer lasers**.
 - Excimer lasers are based on rather strange materials: Compounds of noble gases like **KrF**, or **ArF**. Rather unstable stuff, but emitting at **248 nm** (**KrF**) or **193 nm** (**ArF**). With the **KrF** system, dimensions down to **130 nm** have been realized, but this is already pushing it quite a bit.
 - The **ArF** excimer laser has been used from about **2003**, so it is still in its infancy. It is expected to cover the "**65 nm** node", and possibly also the **45 nm** node.
 - That will be the end. After that, the age of "**EUV**" (extreme ultraviolet) might start, at a wavelength around **12 nm** (its really rather soft **X-rays**). There is no way of having a lens anymore, "optical processing" must then be done with mirrors.
- If we now look at the *lens* issue, we first should realize that high-aperture lenses are generally difficult to make. But the overwhelming issue is to find suitable materials that have a sufficiently large index of refraction at the wavelength considered.
- We already looked at this issue, e.g. in the context of the frequency dependence of the dielectric constant, so we need not repeat the problems encountered here. Check the following links:
 - [General remarks to the frequency dependence of \$\epsilon\$ and \$n\$](#)
 - [Dielectrics and optics](#)
 - [The CaF₂ lens](#)
- Illumination source and lens materials are not the only problems encountered by switching to a smaller wavelength. Of course, there are many others, too.
- To mention just one: The "[pellicles](#)", the thin foils protecting the mask, will turn dark in intense **UV** illumination. Not good, so let's take a better material. Easy fix, but do you know a better material? No? Too bad - since nobody else does either, you missed your change of getting rich quickly.

Numerical Aperture

- ✦ In air, **NA** obviously than has a maximum value of **1**. The best lenses built so far have a **NA** of about **0.8**; but **0.9** is already aimed for
- Keep in mind that what you gain in resolution by increasing **NA**, you loose in the [depth of focus](#). Large **NA** lenses thus only make sense in the context of rather perfect **planarization**.
- ✦ Nevertheless, increasing **NA** even more helps, and there is - in principle - a simple way of doing it: Replace the air between your lens and the wafer with something that has an appreciable index of refraction, e.g. oil.
- "Oil immersion objective lenses" have been used for about a century in conventional optical microscopes; in this way the numerical aperture and thus resolution can be increased in a rather simple way by up to **40%**.
 - But this is far easier said than done. Just consider that the name "[stepper](#)" comes from the fact, that you **step** the wafer (rather rapidly) below the lens. How do you keep you oil in place? And how will the wafer respond to be covered with oil?
 - Well, let's not use oil, let's use high-purity water ($n = 1.437$ at **193 nm**), but that only solves some of many problems and creates some new one (your **CaF₂** lens, for example, will dissolve in water).
 - Nevertheless, "**liquid immersion lithography**" will most likely be the next big fashion in lithography, with the potential to keep microelectronics alive well into the next decade (i.e. after **2010**).

Recticles and Resists

- ✦ What is left is to make the [parameter \$k\$](#) as small as possible, i.e. to pay some attention to **reticles** and **resist**, or, more general, to resolution enhancing techniques.
- There is quite a potential here, "historically" parameter **k** has decreased steadily form about **0.8** in the **1980s** to **0.4** today.
 - While optimizing the resist is critical, it does not introduce new principles, and we will not cover it here.
 - That leaves the reticle and the way it is illuminated. There is quite a bit that can be done, but you must pay the prize of sharp increases in complexity.



- The proper catchwords giving some idea to what is meant are:
 - Off-axis illumination
 - Optical proximity correction (OPC)
 - Phase shift masks (PSM)
 - For the latter two cases the general idea is to have a structure on the reticle that is different from what you want to have projected into the resist on the wafer. If, for example, a sharp corner is "smeared out" to a roundish image, than make the corner look different. The figure gives a rough idea what that means
 - In phase shift masks you add structures that do not only manipulate the amplitude of the light transmitted through the mask, but also the phase.
 - In this way you can produce constructive or destructive interference in the image plane in places where that is helpful to sharpen the image.
- ✦ Of course, all these additional features on the mask must first be computed (not easy), than made (very difficult), and finally tested (exceedingly difficult).
- Testing your mask is essential, that any mistake in the mask will automatically be transferred to the chip and, remember Murphy's law, more likely than not kill the chip.
- ✦ In the grand total a set of masks will quickly cost you up to **2.000.000 €** You must a sell a hell of a lot of chips (at a profit) just to recover that cost
- For customized chips, that are not made by the untold millions, its simply not possible to pay that prize.
 - This drives a large-scale effort to find some better solutions. For mor details (and for the source of some of the data here), refer to "Materials today" from Feb. **2005**.

Multiple Choice Test zu

6.1.2 Silicon Crystal Growth and Wafer Production

Start Multiple Choice

Multiple Choice Test zu

6.2.1 Si Oxide

Start Multiple Choice

Multiple Choice Test zu

6.2.3 Si Oxide and LOCOS Process - Summary

Start Multiple Choice

Multiple Choice Test zu

6.3.3 CVD for Poly-Silicon, Silicon Nitride and Miscellaneous Materials

Start Multiple Choice

Multiple Choice Test zu

6.4.1 Sputter Deposition and Contact Hole Filling

Start Multiple Choice

Multiple Choice Test zu

6.4.2 Ion Implantation

Start Multiple Choice

Multiple Choice Test zu

6.4.3 Miscellaneous Techniques

Start Multiple Choice

Multiple Choice Test zu

6.4.4 Physical Processes for Layer Depositions - Summary

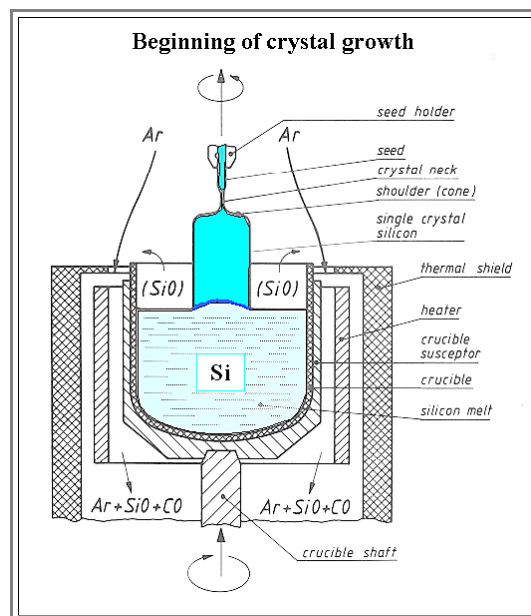
Start Multiple Choice

Czochralski Crystal Growth Process

Illustration

You start growing a "**Czochralski crystal**" by filling a suitable crucible with the material - here hyperpure correctly doped Si pieces obtained by crushing the poly-Si from the [Siemens process](#). Take care to keep impurities out - do it in a clean room - and use hyperpure silica for your crucible.

- Make sure that the inside of the machine is very clean too and that the gas flow - the gas you introduce but also the SiO coming from the molten Si because parts of the crucible dissolve - does not interfere with the growing crystal.
- Dissolve the Si in the crucible and keep its temperature close to the melting point. Since you cannot avoid temperature gradients in the crucible, there will be some **convection** in the liquid Si. You may want to suppress this by big magnetic fields.
- Insert your **seed crystal**, adjust the temperature to "just right", and start withdrawing the seed crystal. For homogeneity, **rotate** the seed crystal and the crucible. Rotation directions and speeds and their development during growth, are closely guarded secrets!
- First pull **rather fast** - the diameter of the growing crystal will decrease to a few mm. This is the "**Dash process**" ensuring that the crystal will be dislocation free even though the seed crystal may contain dislocations.
- Now decrease the growth rate - the crystal diameter will increase - until you have the desired diameter and commence to grow the commercial part of your crystal at a few mm/second.

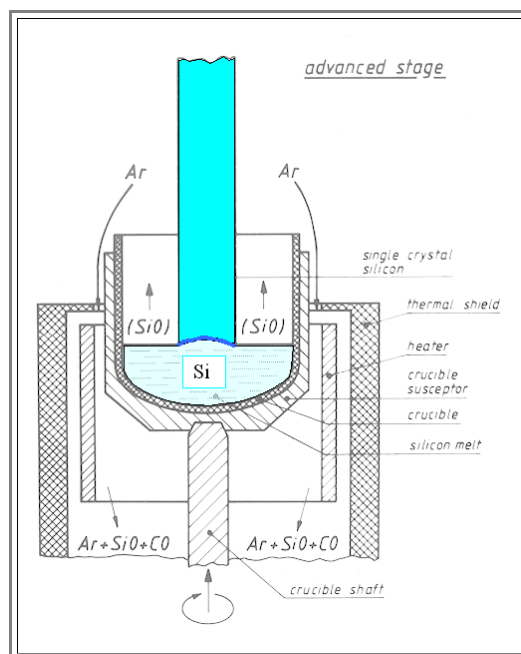


As your crystal grows, the impurity concentration (including the dopants if you do not watch out) will increase in the melt (due to [segregation](#)) and therefore also the percentage incorporated into the crystal. The temperature profile of the whole system will also change - you are now deeper down in the crucible and the crystal cools off a little more slowly. All these factors influence the homogeneity of the crystal.

- The radial and lateral doping level is influenced - it will not stay constant without some special measures
- The concentration of impurities, especially interstitial oxygen, may change. In general, the concentration increases from "head" to "tail".
- Crystal lattice defects still present (essentially agglomerates of the point defects present in thermal equilibrium at high temperatures) may change in size and distribution.

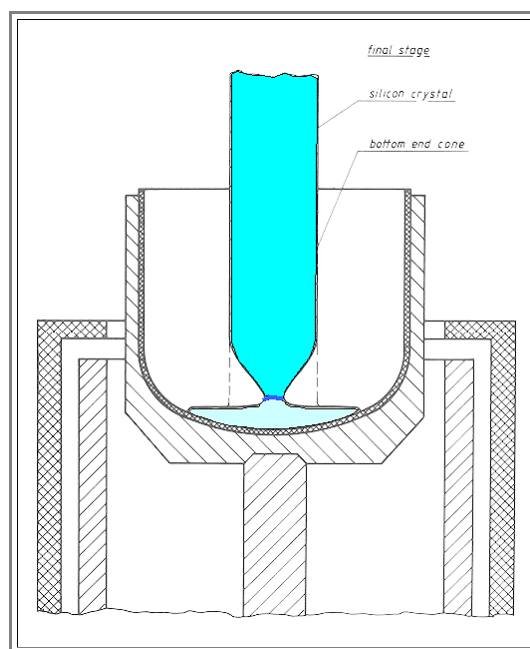
You do not want this - you want a crystal where all these factors are constant - everywhere!

- So you must do something - change the rotation speeds, the temperature, the growth speed - whatever.
- This is where crystal growing becomes an art - and you will not find much literature about this. This is the tricky and secret part: Changing all important parameters continuously so that the crystal is homogeneous!



Now the crystal is nearly finished. You do not want to use up all the **Si**, because the "last drop" contains all the impurities not yet incorporated because of their small segregation coefficients.

- But you cannot simply pull out the crystal after the desired length has been reached. The thermal shock of the rapidly cooling end would introduce large temperature gradients in the crystal which in turn produce stress gradient - plastic deformation (easy in **Si** at high temperatures) will take place and this means dislocation are nucleated and driven into the crystal.
- The dislocation will even run up into the formerly dislocation free part of the crystal, destroying your precious Silicon.
- So you withdraw gradually by just increasing the pulling rate a little bit which will lead to a reduced diameter. The crystal then ends in an "end cone" similar to the "seed cone".



- The [finished product](#) can be seen in a different link.

Silicon Crystal

- Here is a picture of a state-of-the-art **200 mm Si** crystal as they are grown by the thousands for present day (2000) chip manufacture.
- While it does look like an oversized chromium-plated salami, it is a much more sophisticated product (and much more expensive).

Illustration



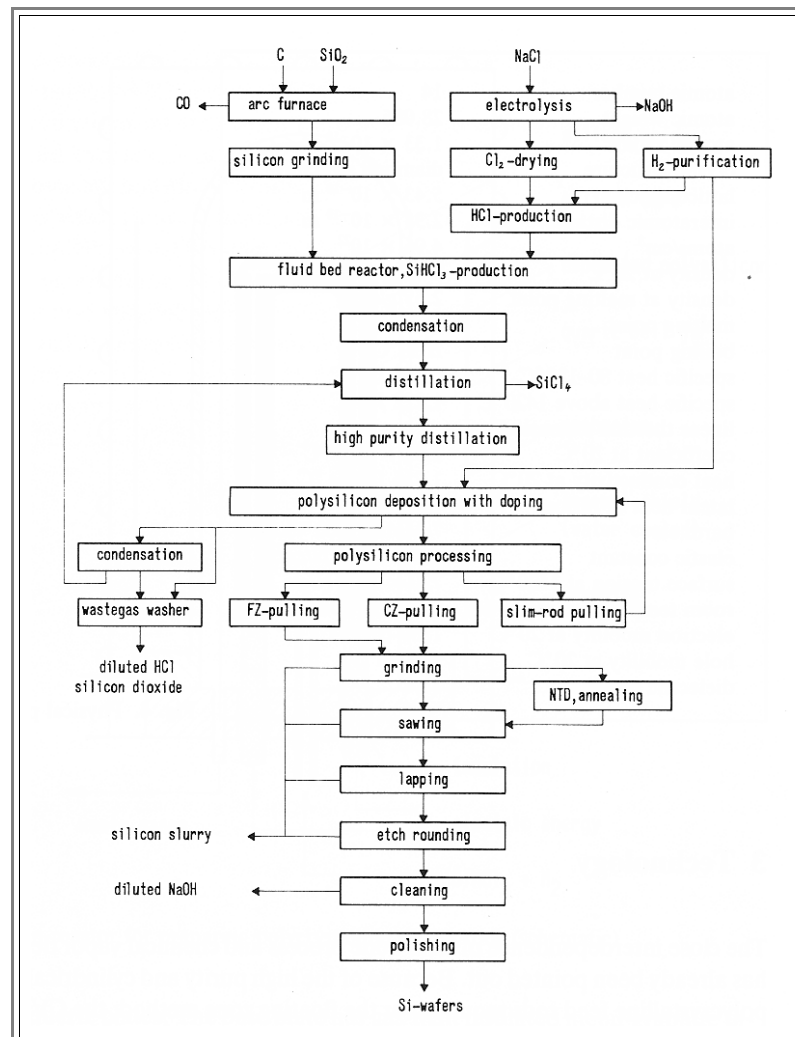
- Note that this huge crystal is hanging on a rather thin **Si** seed crystal (see inset). This seed crystal does not only have to support the weight of the crystal, but also the torque needed to rotate the crystal during its growth.

Complete Wafer Process

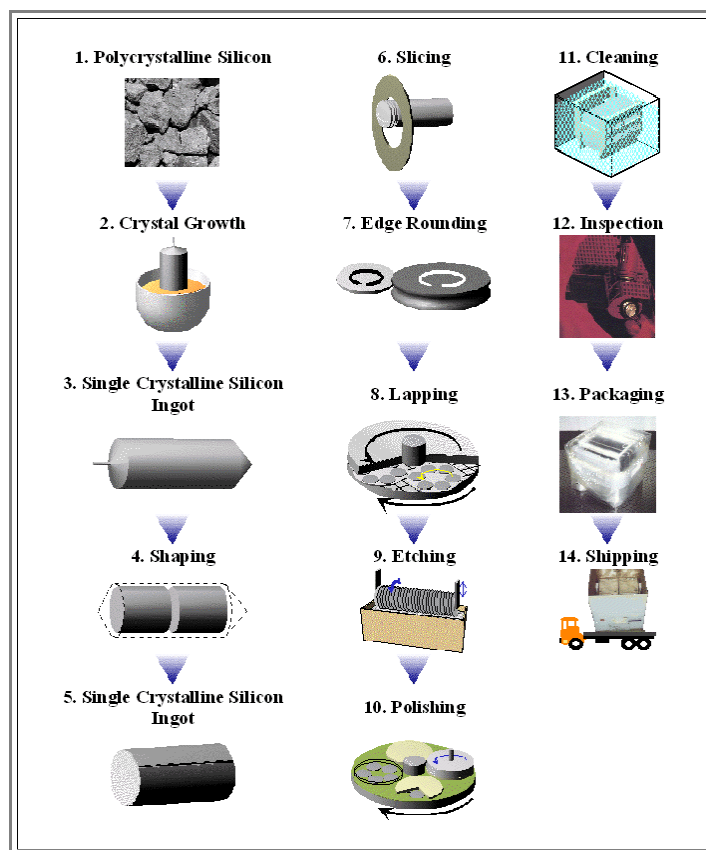
First a schematic process flow graph is given for everything - from sand to wafer.

It includes side processes not covered in the backbone modules which, however, are self explanatory.

Illustration



Next, we have a pictorial representation of just crystal growth and wafer production. Again, it contains some self-explanatory processes not introduced before.



Commercial Poly-Silicon Specifications

Here are the specification for poly-silicon from one of the worlds largest suppliers, **Wacker Siltronic** as they appear in the Internet in Nov. **2000**.

Notice: The "w" or "a" behind the concentration denotes **w**eight or **a**tomistic parts per **m** = million, **b** = billion , **t** = trillion.

Illustration

PolySilicon			
PolySilicon for Crucible Growing			
Chip Size		mm	5 - 45 / 20 - 65 / 20 - 150
Surface			smooth, etched
Surface Metal Concentration	Monitor: Iron	pptw	< 500
Bulk Element Concentration	Donors (P, As, Sb)	ppta	< 150
	Acceptors (B, Al)	ppta	< 50
	Carbon	ppba	< 100
PolySilicon Ingots for Float Zone Growing			
Ingot Length		mm	600 - 1,850
Diameter		mm	90 - 105 / 118 - 135 / 135 - 154
Surface			smooth, etched
Bulk Element Concentration	Donors (P, As, Sb)	ppta	< 300
	Acceptors (B, Al)	ppta	< 100
	Carbon	ppba	< 200
PolySilicon Ingots for Crucible Growing			
Ingot Length		mm	320 - 980
Diameter		mm	90 - 115 / 115 - 135
Surface			smooth, etched
Bulk Element Concentration	Donors (P, As, Sb)	ppta	< 300
	Acceptors (B, Al)	ppta	< 100
	Carbon	ppba	< 200
Solar Grade PolySilicon for Crucible Growing/Casting			
Chip Size		mm	0 - 15 / 5 - 160
Bulk Element Concentration	Donors (P, As, Sb)	ppta	< 300
	Acceptors (B, Al)	ppta	< 100
	Carbon	ppba	< 200
Cleaning			none

Here some productio information:

According to "Solid State Technology" July **2005**, the productio numbers are as follows:

- Total production 2005: **26.000.000 kg**; about **2/3** for miocroelectronics, **1/3** for photovoltaics.
- Expected production 2006: **29.000.000 kg**.

Expected shortfalls:

- 2005: **4.000.000 kg**
- 2006: **6.000.000 kg**.
- 2007: **12.000.000 kg**.
- 2008: **20.000 000 kg**.

The expected shortfalls result to a large extent from a growthrate of **30 %** for photovoltaics and from technical and financial difficulties to crank up püroduction at a high reate. However, alternative processes solar **Si** production are expected to come on-line in 2006.

Commercial Wafer Specifications

Illustration

Here are the specification for Si wafers from one of the worlds top companies, **Wacker Siltronic**, as they appear in the Internet in Nov. **2000**.

- Notice:** Concentrations here are in cm^3 . The conversion to parts per .. is simple:
The atomic density of Si is $4.96 \cdot 10^{22} \text{ cm}^{-3}$ or about $5 \cdot 10^{22} \text{ cm}^{-3}$.
1 ppm thus corresponds to $5 \cdot 10^{16} \text{ cm}^{-3}$.
- The lowest concentration given in the table (look for it) is $5 \cdot 10^{10} \text{ cm}^{-3}$; it corresponds to **1 ppt** or 10^{-12} .
- Surface concentrations **[S]** (given in cm^{-2}) are converted to volume concentrations **[V]** by
[S] = [V] / a with **a** = lattice constant (= **0,5431 nm**) or, more precise for single crystals, distance between the crystallographic planes. With **a** \approx **0,5 nm** = $5 \cdot 10^{-8} \text{ cm}$, we have
[V] = $5 \cdot 10^{16} \text{ cm}^{-3}$ = 1 ppm corresponds to **S = 10^8 cm^{-2}** .
- Many specifications relate to the "flatness" of the wafers and the perfection of the surface; the abbreviations used are
LLS (sometimes also abbreviated **LPDs**): **Localized Light Scattering Defect**; this relates to a detection method of **sub- μm** size surface imperfections (resulting from bulk microdefects)
SFQR: **Site flatness quality requirements** (??): Definitely a measure of flatness in a region comparable to the size of a single chip
(The rest: Who knows?)
- Here is a link with precise defininitioons of geometricla parameters:
http://www.freiberger.com/english/products/geom_parameters.php

Polished & Epitaxial Wafers for IC Applications

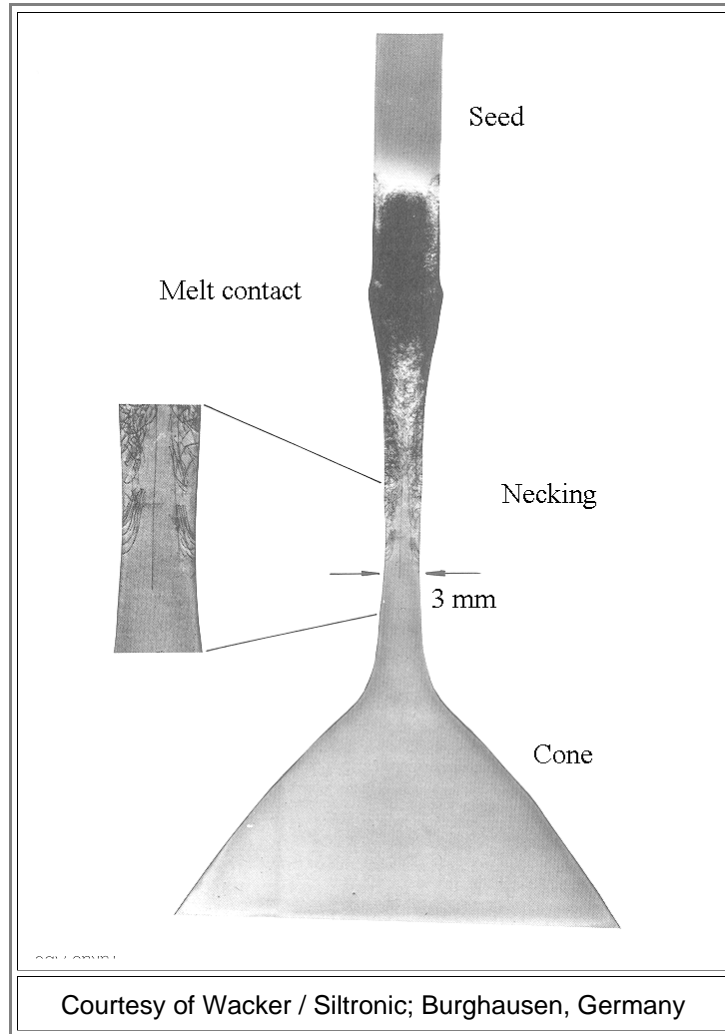
Crystal / Bulk			300mm	200mm	150mm	125mm	100mm	
Growth Technique *)			CZ	CZ	CZ	CZ	CZ	
Orientation			1-0-0	1-0-0	1-0-0 / 1-1-1	1-0-0 / 1-1-1	1-0-0 / 1-1-1	
Orientation Tolerance		degree	± 0.2	± 0.2	± 0.5	± 0.5	± 0.5	
Off Orientation		degree	0	0 - 4	0 - 4	0 - 4	0 - 4	
Dopant			Boron / Phosphorus	Boron / Phosphorus	Boron / Phosphorus	Boron / Phosphorus	Boron / Phosphorus	
Resistivity Target Range	pol prime - Boron	Ohmcm	0.5 - 50	0.5 - 50	0.5 - 50	0.5 - 50	0.5 - 50	
	pol prime - Phosphorous	Ohmcm	1.0 - 50	1.0 - 50	1.0 - 50	1.0 - 50	1.0 - 50	
	epi substrate - Boron	Ohmcm	0.006 - 50	0.006 - 50	0.006 - 50	0.006 - 50	0.006 - 50	
Radial Resistivity Variation	Boron typical	1-0-0 / 1-1-1	%	< 10	< 5 / < 6	< 6 / < 10	< 8 / < 9	
	Phosph. typical	1-0-0 / 1-1-1	%	< 15	< 15	< 12 / < 25	< 12 / < 25	< 15 / < 25
	pol prime - Boron	1-0-0	4.8 - 7.8 × 10 ¹¹ ± 0.5	5 - 7.8 × 10 ¹¹ ± 0.5	5.8 - 8.9 × 10 ¹¹ (± 0.6 - 0.8)	5.8 - 8.9 × 10 ¹¹ (± 0.5 - 1.0)	5.8 - 8.9 × 10 ¹¹ (± 0.8 - 1.2)	
Oxygen Target Range ± Tol.		1-1-1	NA	NA	5.8 - 8.9 × 10 ¹¹ (± 0.7 - 1.0)	6.2 - 8.9 × 10 ¹¹ (± 0.5 - 1.0)	5.9 - 8.9 × 10 ¹¹ (± 0.8 - 1.5)	
	pol prime - Phosph.	1-0-0	ASTM F121-83	4.8 - 7.8 × 10 ¹¹ ± 0.5	6 - 7.5 × 10 ¹¹ ± 0.5	5.8 - 8.9 × 10 ¹¹ (± 0.6 - 0.8)	5.8 - 8.9 × 10 ¹¹ (± 0.5 - 1.0)	5.8 - 8.9 × 10 ¹¹ (± 0.8 - 1.2)
		1-1-1	NA	NA	5.8 - 8.9 × 10 ¹¹ (± 0.7 - 1.0)	6.2 - 8.9 × 10 ¹¹ (± 0.5 - 1.0)	5.9 - 8.9 × 10 ¹¹ (± 0.8 - 1.5)	
Radial Oxygen Variation	typical	%	< 10	< 5	< 6	< 6	< 5 - 10	
Bulk Metal Concentration	Fe	at cm ⁻³	≤ 5.0 × 10 ¹⁰	≤ 5.0 × 10 ¹⁰	≤ 1.0 × 10 ¹¹	≤ 1.0 × 10 ¹¹	≤ 1.0 × 10 ¹¹	
Bulk Carbon Concentration	measured on wafer	at cm ⁻³	≤ 2.0 × 10 ¹¹	≤ 2.0 × 10 ¹¹	≤ 2.0 × 10 ¹¹	≤ 2.5 × 10 ¹¹	≤ 2.5 × 10 ¹¹	

Polished Wafers / Substrates			300mm	200mm	150mm	125mm	100mm		
Surface Metals	Cu / Cr / Fe / Ni	at cm ⁻²	≤ 1.0 × 10 ¹¹	≤ 2.5 × 10 ¹¹	≤ 5.0 × 10 ¹¹	≤ 5.0 × 10 ¹¹	≤ 5.0 × 10 ¹¹		
	Al / Zn / K / Na / Ca	at cm ⁻²	≤ 5.0 × 10 ¹¹	≤ 1.0 × 10 ¹¹	≤ 2.0 × 10 ¹¹	≤ 2.0 × 10 ¹¹	≤ 2.0 × 10 ¹¹		
	size	µm	> 0.2	> 0.16	> 0.12	> 0.3	> 0.12	> 0.3	
LLSs (Frontside) *)	pol prime	# per wafer	< 30	< 40-300	< 200-10 ¹	< 15-35	< 20-120	< 70-600	
	UltraFlat (150 mm)	# per wafer	NA	NA	NA	NA	< 5	< 10	< 150
	monitor	# per wafer	< 30	< 60	< 100	< 15	< 20-65	< 130-700	NA
Diameter Tolerance		mm	± 0.2	± 0.2	± 0.2	± 0.2	± 0.2	± 0.2	
Warp	polished - without layer	µm	< 50	< 20	< 30	< 30	< 30	< 30	
Wafer / Substrate Thickness	Standards	µm	775	725	375 / 525 / 625 / 675	375 / 525 / 625	300 / 375 / 525		
Thickness Tolerance		µm	± 25	± 15	± 15	± 15	± 15	± 15	
GBIR = TTV (Std UltraFlat*)		µm	< 4	< 3.5	< 5.0	< 2.5	< 5.0	< 5.0	
GFLR = TIR (Std UltraFlat*)		µm	NA	NA	< 2.0	< 1.2	< 2.0	< 2.0	
Local Flatness *)	SFQR / STIRmax, s.b.f.	µm	< 0.25	< 0.25	< 0.5	< 0.3	NA	NA	
	SFQD / SFPD, s.b.f.	µm	< 0.18	< 0.18	< 0.3	< 0.2	NA	NA	
	SBIR / STIRmax, b.r.	µm	NA	< 0.7	< 1.0	< 0.6	< 1.0	< 1.0	
Standard Site Size		mm ²	25 x 25	25 x 25	15 x 15	15 x 15	15 x 15		

Necking

Here is an X-ray topograph of the first part of crystal growth

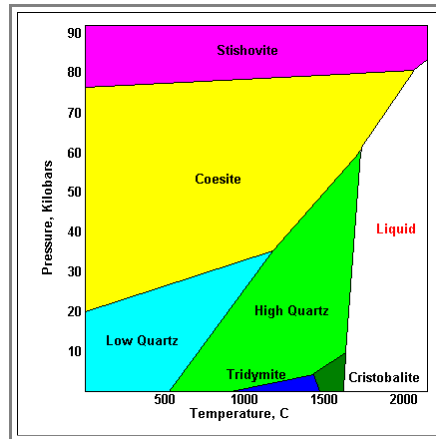
- An X-ray topograph is similar to a transmission electron microscope image - it shows the interior of the sample and dislocations are visible as dark lines.



Forms of SiO₂

Here is a **phase diagram** of **quartz**; unfortunately I do not know the source

What we have at room temperature and standard pressure is "low quartz" or α - quartz

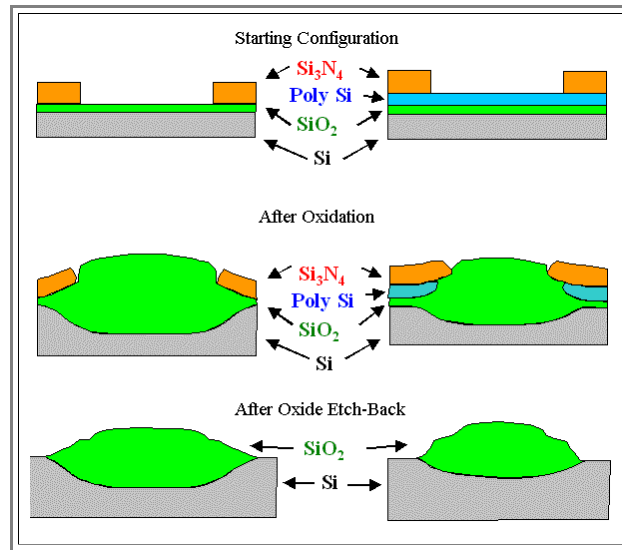


Illustration

LOCOS with Sacrificial Poly-Silicon

Illustration

Here is a comparison of the [LOCOS process](#) with and without some sacrificial poly-Si:

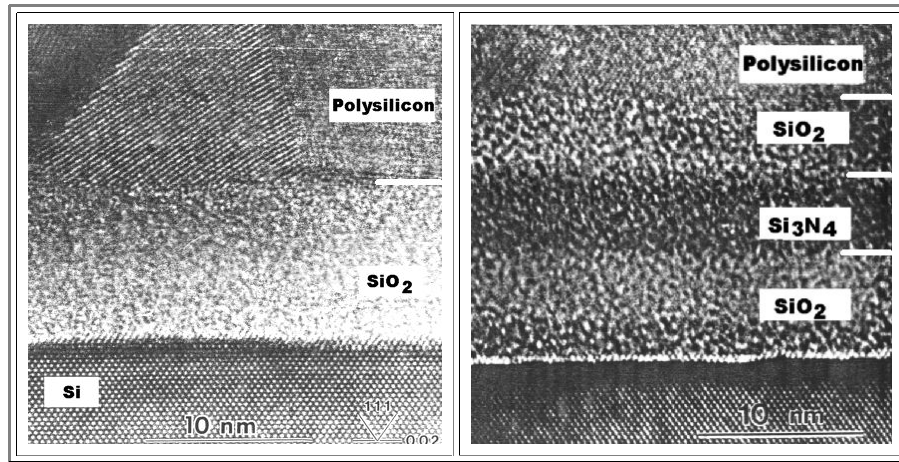


- Provided, some of the oxide is removed by an "etch-back" process, the lateral extension can be kept somewhat smaller than in the conventional case.

While this makes things quite complicated, the final versions of the **LOCOS** process were even more complicated.

HRTEM of Gateoxide and ONO



Illustration



Furnaces

Illustration

- Furnaces for thermal oxidation, but also simply for annealing or for **CDV** processes are complicated and expensive pieces of equipments.
- While horizontal furnaces dominated the scene for **150 mm** wafers or smaller, with **200 mm** wafers a switch to vertical furnaces took place
- Below two pictures showing a horizontal and a vertical furnace for **200 mm** or **300 mm** wafer, respectively. Of course, you don't see much; nevertheless, they are big pieces of equipment.

	
<p>Horizontal oxidation furnace; three tubes</p>	<p>ASM A412 300mm twin vertical furnace system for high temperature atmospheric pressure oxidation and LPCVD processing of polysilicon</p>

- Here is a comparison between horizontal and vertical furnaces, taken straight from the homepage of a major furnace manufacturer.


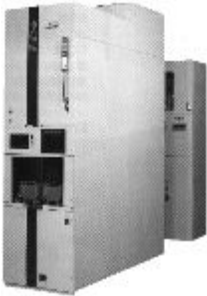

Comparison Vertical vs. Horizontal Furnaces

Koyo Thermo Systems manufactures numerous versions of [vertical furnaces](#) and [horizontal furnaces](#) with full automation or manual loading. Smaller versions for pilot and laboratory applications are available.

We are frequently asked for the differences of vertical furnaces and horizontal furnace and for a justification of the higher price of vertical systems. Therefore we worked out the following table sheet. We hope that this overview, this comparison sheet will help you to make a proper planning.

We compare in this sheet a horizontal furnace with a vertical furnace for mass production (VF5300) and with our small vertical furnace VF1000, which is a good alternative to horizontal systems for small companies as well as for research institutes.

Feature	3- or 4-Tube Horizontal Furnace e.g. Model 206	Vertical Furnace for Production e.g. VF5300	Vertical furnace for RD e.g. VF1000

			
Heat up speed	slower	faster	faster
Quartz boat loading time	slower	faster	faster
Cool down speed	equal	equal	equal
Max. temperature using KLL's advanced LGO heaters	1100°C	1250°C	1250°C
Temperature uniformity	lower	higher	higher
Temperature interference between the tubes	1 - 2 °C	none	none
Oxygen concentration tube center (end open)	16%	0,1%	0,1%
Oxygen concentration tube end (end open)	high	500ppm	500ppm
Oxygen concentration tube center (end closed)	0,1%	300ppm	300ppm
Oxygen concentration tube end (end closed)	10 - 30 ppm	10ppm	10ppm
Air tight process chamber (atmospheric process)	no	yes	yes
HCl leak free	no	yes	yes
Cross contamination	possible	not possible	not possible
Process independence	not completely	yes	yes
Particle data	worse	very good	better
Flexibility: Run mixed diameter of wafers in one run	possible	not possible	not possible
Flexibility: Run different diameter of wafers run to run	possible	not possible	possible
Flexibility: Range of wafer diameters that can be processed	3" - 6" (8")	4" - 300mm	3" - 300mm
Stock wafer cassettes	no	yes	no
Automation level	lower	very high	higher
Thickness uniformity wet oxidation 10nm, 8" wafer	no data	± 0.9 %	± 0.9 %
Thickness uniformity dry oxidation 20nm, 8" wafer	± 2.4 %	± 1.2 %	± 1.2 %
Thickness uniformity poly-Silicon 400nm, 8" wafer	± 2.0 %	± 1.0 %	± 1.0 %
Thickness uniformity Nitride 100nm, 8" wafer	± 2.5 %	± 1.5 %	± 1.5 %
Capacity	150 wafer	100 - 150 wafer	25 wafers
Power consumption	high	lower	very low

Maintenance independence	no	yes	yes
Maintenance work, necessary	higher	lower	very low
Footprint / tube	2.6 - 3.4 m ² (partially cleanroom)	3.0 m ² (grey room)	1.5 m ² (grey room)
Price	low	high	low

Nowadays, mass production of semiconductor chips happens with silicon wafers with 200mm or 300mm diameter. Vertical furnaces are used almost exclusively. Only in older factories, which still use smaller wafer diameters, horizontal furnaces are still common. For wafer diameters until 150 mm the performance of such systems is in many cases still good enough to fulfil the customer requests. However, the advantages of vertical systems are already evident for this wafer size.

The result of this situation on the oven market was, that almost all large furnace manufacturers stopped the further development of horizontal furnaces. Development work is done today almost only for vertical systems. Therefore vertical furnaces are superior to horizontal ones not only for physical reasons, but also because they are the more modern production tool. Their performance is much higher.

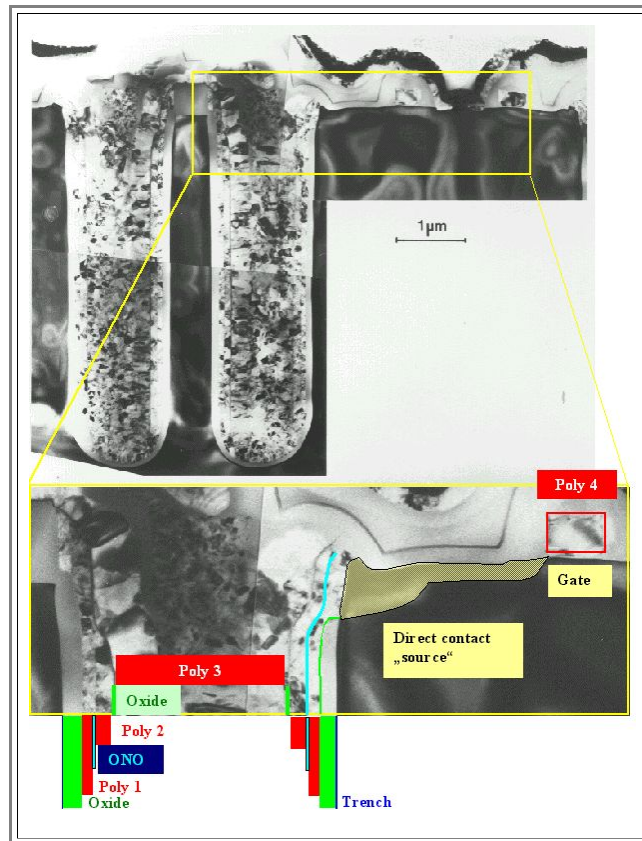
A main issue is furnace automation. Automation for horizontal furnaces means mainly the installation of an elevator system for the loading of the boat on the cantilever arm. The loading area is open to the clean room. Vertical furnaces however are closed system with clean room class 1 inside. The loading happens fully automatically from the cassette by advanced robot systems.

Other technical advantages of the vertical furnaces are the better gas tight sealing of the furnace tube, as well as several options, available only on vertical furnaces like improved temperature uniformity by boat rotation or nitrogen load lock chamber.

Poly-Silicon

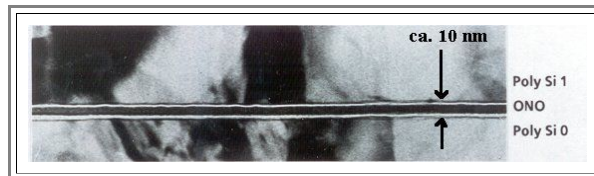
Illustration

Here is a large size rendering of the **TEM** picture of a **16 Mbit DRAM** memory cell. A [drawing of this cross section](#) can be found in the link.



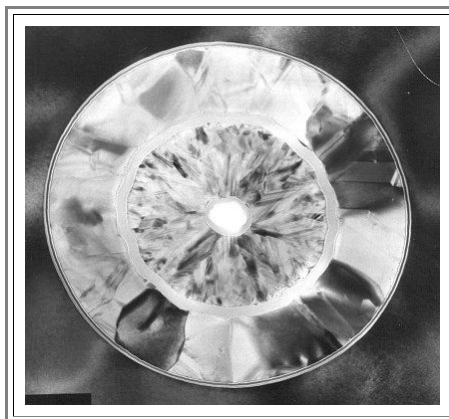
.Many details are not visible in this low magnification picture. Below we see the ONO layer between the two poly Si layers at high magnification

- The two oxide layers appear white; the nitride layer dark



Next, we look at a trench capacitor "*from above*" and not in *cross section*.

- Shown is the "simple" capacitor from the **4 Mbit DRAM** generation. **ONO** is used as dielectric; and the substrate **Si** served as one electrode.
- Again the poly electrode was oxidized for insulation and the trench filled with poly. The smaller (and dendritic) grain structure indicates that a large deposition rate (at somewhat higher pressure) was used; and a little hole remained unfilled in this case.

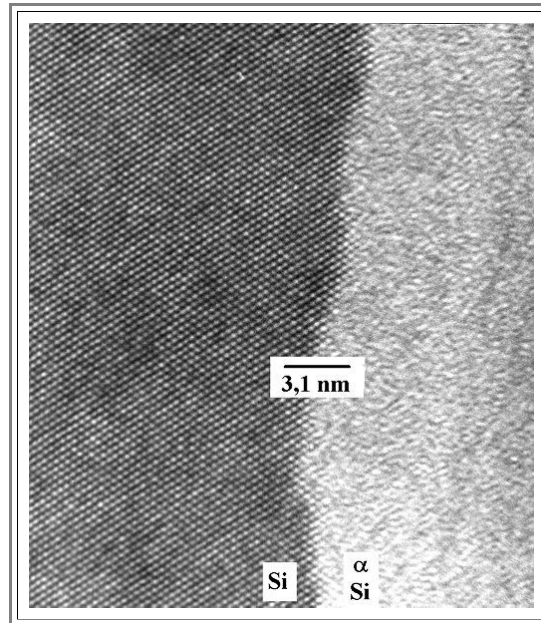


Amorphization after Ion Implantation

Illustration

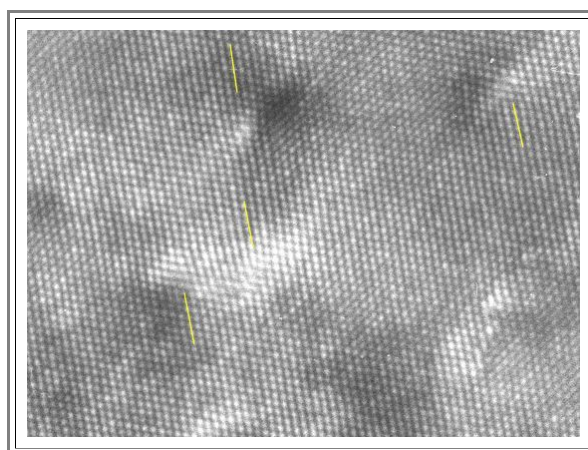
The picture below shows the boundary between crystalline **Si**, and **Si** that has been rendered *amorphous* by an ion implantation (the ion beam came from the right). The picture is slightly remarkable because it was the first [high resolution transmission electron microscopy \(HRTEM\)](#) picture ever taken from damage after ion implantation.

- There is indeed an amorphous **Si** (α -**Si**) layer. (Think a minute how you could ascertain that without transmission electron microscopy).
- There is a pretty abrupt, if somewhat wavy boundary between the amorphous and the crystalline **Si**.
- There seems to be little disturbance in the **Si** lattice - it looks pretty perfect.



Quick glances at **HRTEM** pictures may be deceiving, however.

- A lot of point defects may be contained in the lattice - they would not clearly show in this picture
- Looking a bit more closely at some greater depth (to the left of the above picture), a high density of dislocations is found. An example is shown below; the ending lattice planes are indicated with yellow lines.
- Note that not all dislocations will show up in this kind of imaging mode.



Loosing Large Amounts of Money with Wet Chemistry

Illustration

- Usually, the big disasters in chip production are kept quite confidential - you don't want to admit that you have problems, and you don't want to help your competitors to avoid these problems.
 - However, as time goes by, confidentiality is no longer necessary. But people have other problems now, so still nothing is published.
 - Well, here is the exception: The short story of a major disaster in the early production of the **1 Mbit DRAM** in the new Siemens factory in Regensburg, as it happened back in about **1985**.
- Managerial wisdom had decided not to trust the research and development team with the development of the **1 Mbit DRAM** (Siemens (and Phillips) then were about **1** year behind the leading Japanese and had started a race to catch up), but to take a license from Toshiba, the top memory producer then.
 - Not that we liked it. But the new factory was dutifully converted to the Toshiba process - all the equipment, all materials, everything whatsoever, was identical to what Toshiba had and did - with good yields of functioning devices.
 - But the German factory produced exclusively junk - the memory chips didn't work. Nobody, including the Toshiba engineers, had the faintest idea why.
 - This went on for almost **6** months - at losses of about **5 - 10** million marks a month. Then the problem was found and solved. First empirically, then by understanding what happened.
- Since it was evident that something **we** did must be different from what **Toshiba** did, the search focussed on the few differences that were unavoidable for some reason or other.
 - The culprit that was finally identified, was an extremely simple chemical: **H₂O₂**; used in the mixtures for cleaning the wafers.
 - In contrast to practically all other chemicals, the **H₂O₂** was not bought in Japan from the source Toshiba used, but from a German company because it simply would not have survived months of traveling aboard a ship. **H₂O₂** always decays into water and **O₂** in the course of weeks, and since it is a slightly dangerous chemical, the airlines refused to transport it.
 - So there was no choice but to buy it in Germany - and of course the German **H₂O₂** was carefully checked for cleanliness (it was actually cleaner than the Japanese stuff).
- Somebody finally convinced an airline (Alitalia) to fly in a barrel of the Japanese stuff - and a miracle happened and good chips were produced using it.
- What has happened? Nobody knew, but who cares if it works? Well the research oriented guys do care, and in due course the mystery was unraveled.
 - As it turned out, **H₂O₂** always contains some **stabilizer**, and this is neither displayed on the label nor do the producers tell you what it is. The stabilizer is needed to keep the remaining traces of metal ions, that are still present - even in ultrapure **H₂O₂** - complexed (i.e. surrounded by the stabilizer molecules); in particular **Fe⁺⁺**.
 - And this complexation is necessary because "naked" metal ions would catalyze the decay of **H₂O₂** into water and oxygen - the "shelf life" of your chemical would be very short without a stabilizer.
 - Now Siemens, as most other western producers, used some variant of the classical "**RCA**" cleaning procedure which is always acidic, i.e. it works at **pH** values **<< 7**.
 - The Japanese, however, had invented a new alkaline cleaning procedure, relying heavily on "Choline", a simple organic leach, i.e. they worked at pH values **>> 7**.
- As it turned out, the stabilizer in the Japanese **H₂O₂** worked in an alkaline environment, while the German stabilizer did not.
 - This was purely accidental, neither the Japanese, nor the Germans, nor anybody else, had ever to worry (or even knew) about the stability of **H₂O₂** stabilizers.
- As a consequence, whenever an alkaline cleaning was carried out with the German **H₂O₂**, Fe was no longer complexed and some of it was deposited on the **Si** substrates.
 - This must be expected to happen, because **Si** is less noble than **Fe** with respect to its electrochemical potential. We are talking tiny amounts of deposited **Fe** here, far less of what is still contained in ultra-hyperpure chemicals.
 - The iron deposited in this way would diffuse into the **Si** as soon as it was heated. This did not do much damage, and that was why every measurable parameter always looked quite good during processing - only at the end the chips started to deteriorate.
 - The reason for this was that at every heating cycle, **Fe** was diffusing around a bit more, until eventually small precipitates formed (**needle-like FeSi₂**)
- And these precipitates killed some gate and capacitor oxides - and since it needs only **one** dead transistor (out of about **1,5** million) to kill a chip, the yield was practically non-existent.

- There is just no way you can anticipate that. And the detective work in this case was complicated because the effect (dead transistor) was not traceable to the reason (incorporation of iron), because first measurable deviations from expected behavior occurred many process steps *after* the original cause.
- Many process lines could tell similar stories. From what one hears or suspects, one rule that might be good to know evolves: The really big disasters in chip manufacture are more likely to have their roots in humble wet chemistry than in the sophisticated processes everyone talks about.