LS Scienza dei Materiali - a.a. 2006/07

Fisica delle Nanotecnologie – part 3

Version 5a, Oct 2006

Francesco Fuso, tel 0502214305, 0502214293 - fuso@df.unipi.it
http://www.df.unipi.it/~fuso/dida

Tecnologie convenzionali nell'approccio top-down; II: litografia ottica e suoi limiti

20/10/2006 – 14.30-16.30 – room T1
27/10/2006 – 14.30-16.30 – room T1

Introduction to the topic I

Within the term **lithography** we often include all processes needed to define a lateral pattern onto a substrate (or a multilayered structure)

In an oversimplified picture:
We use an **(material or immaterial)** “ink” to define a pattern;
We can either define a complex pattern all at once **(parallel process),** or draw it “pixel by pixel” **(serial, or scanning, process)**;
The ink is used to modify *locally* the substrate surface **(impression)**;
If required, the substrate can be previously prepared so to be sensitive to the ink **(photoresist or resist-assisted** lithography)
An **etching** process is used to locally remove the material (either in impressed or non impressed regions, i.e., for **negative or positive** lithography), thus to transfer the defined pattern to the substrate

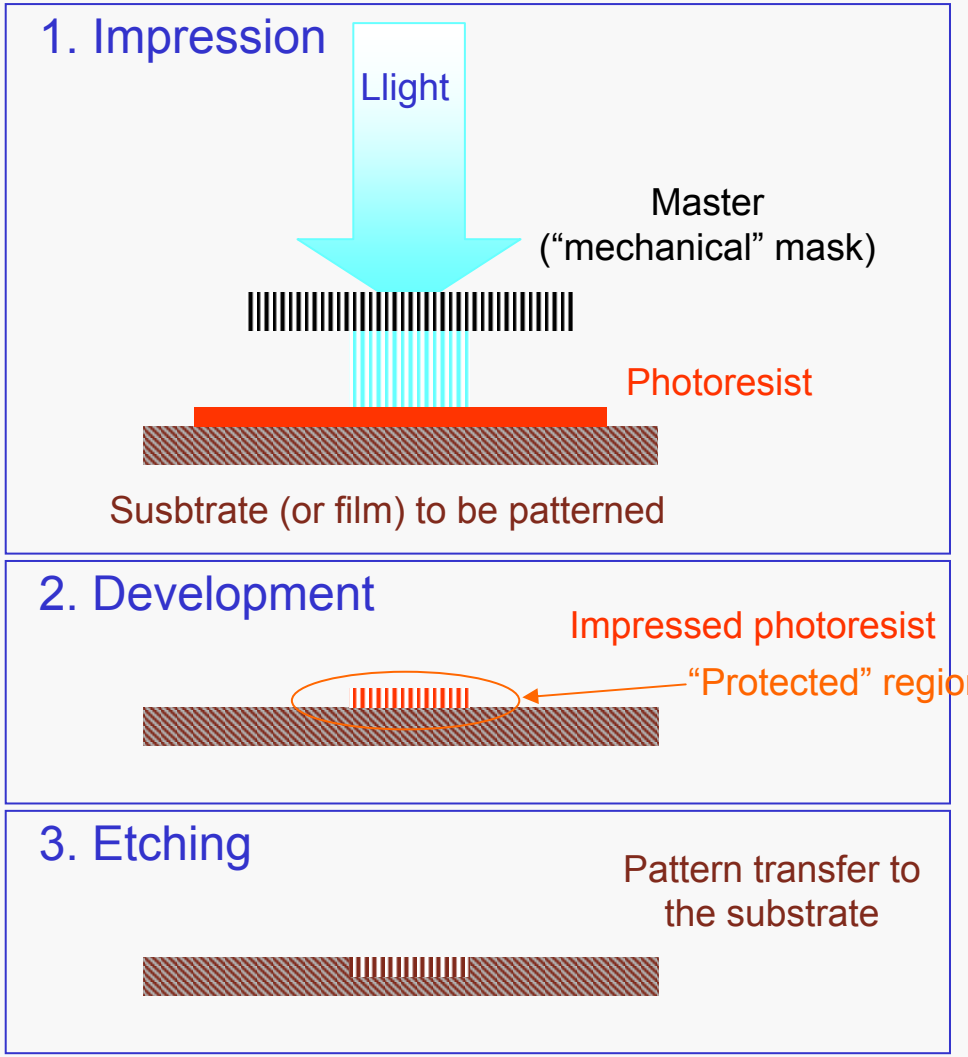
Note: sometimes, instead of etching, **direct deposition** can be used (*as we will see, e.g., in atom lithography*), leading to a ***bottoms-up*** process

Thus, lithography comprises of several aspects, including, e.g.,:
-Methods for “ink” patterning;
-Techniques for impression;
-Etching processes

Introduction to the topic II

Optical lithography is the most common (and most simple) method in microelectronics

Simplified picture of a subtractive process



Essential components:

Light (to make impression)

Mask (to produce the pattern)

Photoresist (to be impressed)

Development and etching (to transfer the pattern)

Inherent advantages:
- *parallel operation* (large-scale impressions in a single run)
- definition of arbitrary patterns

Introduction to the topic III

Fabrication of Sub-45-nm Structures for the Next Generation of Devices: A Lot of Effort for a Little Device

The silicon VLSI* juggernaut is approaching an interesting transition period. For the last four decades, progress in miniaturizing microprocessors has been on an exponential curve, with feature sizes being reduced by a factor of two roughly every 18 months. The result has been not only a tremendous increase in computational power, but also a substantial reduction in the cost of computing, as measured by cost per function, of nearly 30% annually.[1] Today, we fabricate transistors at literally one-billionth the cost of 1950. This kind of progress is almost without precedent in human history. The printing press did not make books a billion times cheaper. Modern agribusiness did not make food a billion times cheaper. The loom did not make clothes a billion times cheaper. But VLSI did make transistors a billion times cheaper. As a result, the world as we know it has been remarkably changed over the last 50 years. The challenge now facing us is how to stay on that exponential curve, known popularly as Moore's law, for the next 50 years. It will be quite a trick.

*VLSI stands for very-large-scale integration, referring to electronic chips with 10^5–10^6 transistors onboard.

There are three basic challenges to maintaining Moore's law: processes and materials that allow continued reductions in size, device physics at the nanoscale, and economics. The basic working assumption of the standard CMOS (complementary metal oxide semiconductor) transistor is that matter is continuous and uniform, and devices can be continuously shrunk without having to worry about the fundamental graininess of matter, the atoms. With device features measured in micrometers, this has been a reasonable assumption for four decades. As device structures shrink to well below 100 nm, this basic working assumption will completely break down in 10–20 years. It will be impossible to build a 10-nm device that works anything like the current CMOS transistor. It will be a nanodevice with new operating principles and physics. It will also require a completely new way of making it.

That is not to say that the challenges facing VLSI fabrication based on the current CMOS design for the next 10 years will be trivial. Far from it. The international roadmap for semiconductors predicts that devices at the 45-nm node will be in production by 2010, with 32-nm devices to follow by 2013.[1] Manufacturing of these devices will require new approaches to patterning, deposition, and etching as current techniques run into fundamental roadblocks. The winning technical solution must also be an economical one: the cost of scaling to smaller dimensions cannot raise the cost of computing, storage, or memory. Moreover, any new process technology for CMOS devices will need to be seamlessly integrated into the existing Si wafer fabrication infrastructure to be broadly adapted.

New technologies to appear in the mid-term

Present technologies to be pushed at their limit in the short-term (one decade?)

MRS Bull., vol. 30 (dec. 2005)

Outlook

1. Optical lithography as the most common technique in microelectronics
   A. Basic processes;
   B. Photoresists and masks

2. Projection lithography and optical microscopy: the (unavoidable) problem of diffraction and space resolution

3. A few words on confocal microscopy

4. Strategies to improve space resolution in lithography:
   i. “optical” methods;
   ii. “chemical” methods:
   iii. decreasing the wavelength…

5. A few words on X-ray lithography and its technological limitations

1.A. Basic processes in optical lithography

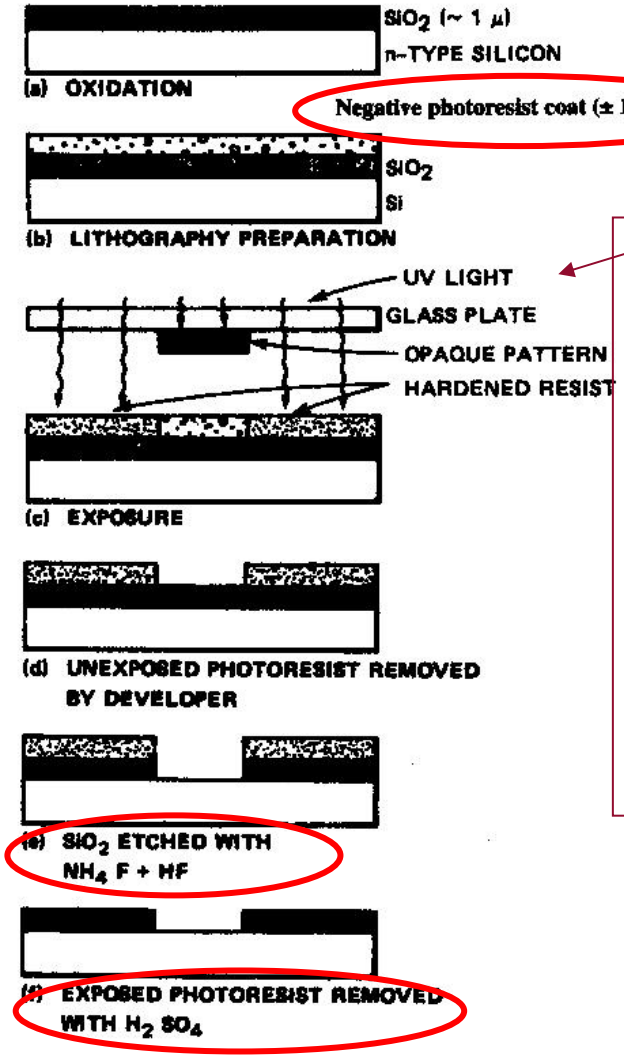


Figure 1.22 Basic IC process steps on an oxidized Si wafer; photolithography (with a negative-tone resist), including exposure, development, oxide etching, and resist stripping. (From Brodie. I. and J. J. Muray, The Physics of Microfabrication, Plenum Press, New York, 1991. With permission.)

Da M. Madou, Fundamentals of microfab., CRC (1997)

Complex sequence of steps (even in very basic processes)

Negative and positive resists exist

Space definition *always* affected by "geometrical" issues (e.g., in the impression or etching stage)

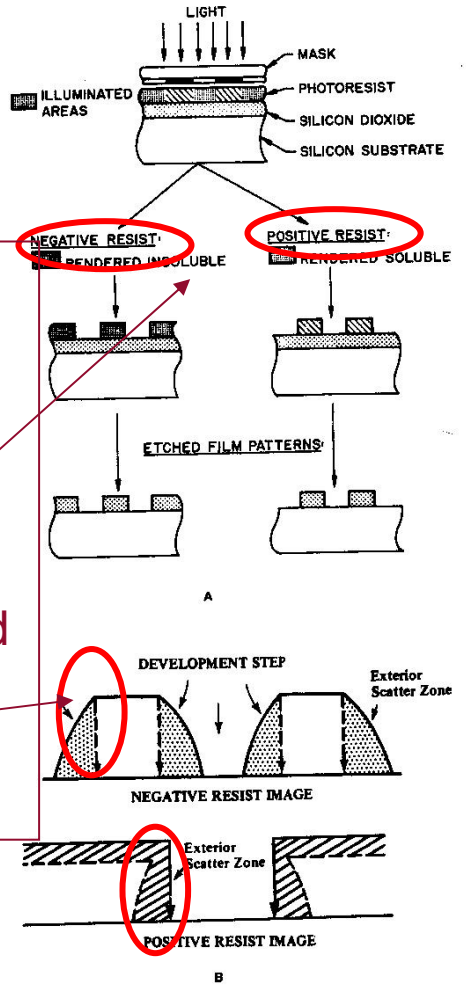
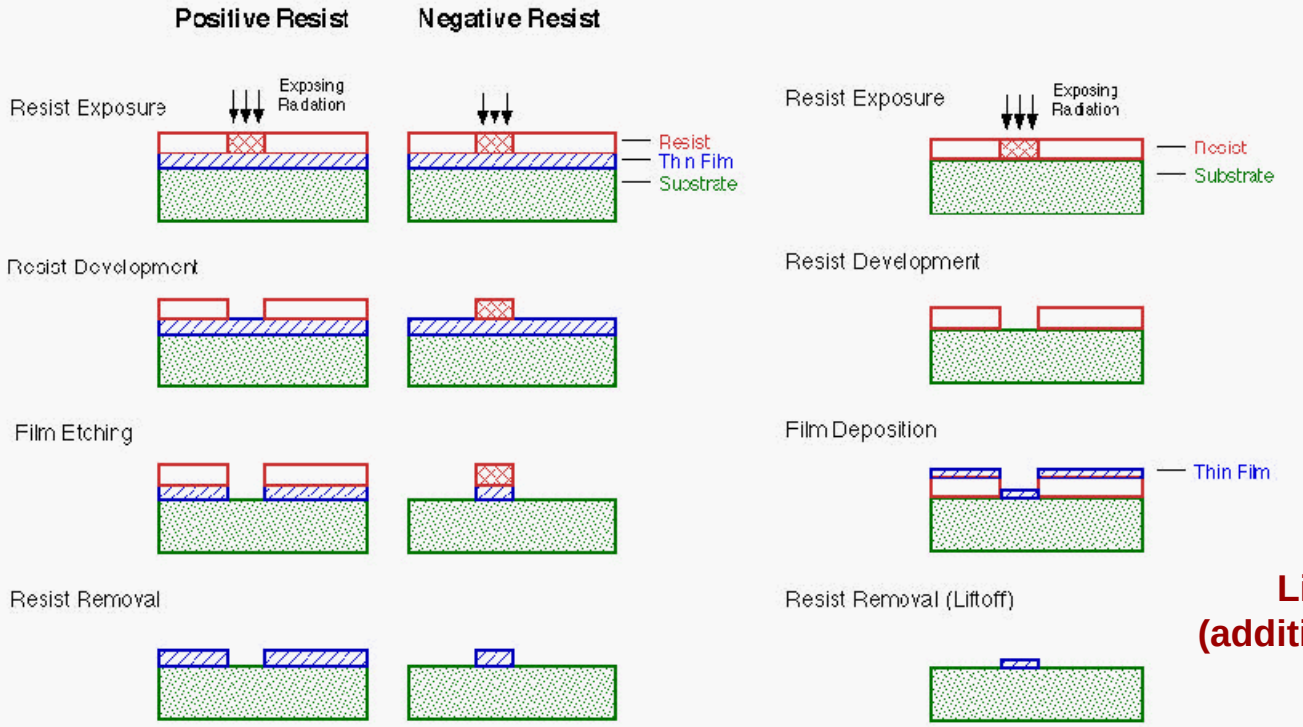


Figure 1.1 Positive and negative resist exposure, development, and edge-scattered radiation. (A) Positive and negative resists, exposure, and development. Positive resists develop in the exposed region and usually remain soluble for lift-off. Negative resists remain in the exposed region but are insoluble and not suitable for lift-off (see text). (B) Edge-scattered radiation profile for negative and positive resists. Time-independent development of cross-linked negative resist fails to remove light scatter zone. Development of positive resist rapidly removes exposed region and can be quenched to inhibit removal of lateral scattered exposed resist region. (From Brodie, I. and J. J. Muray, The Physics of Microfabrication, Plenum Press, New York, 1982. With permission.)

Example of process flexibility in lithography

Two type of pattern transfer(subtractive or additive)



Liftoff (additive litho.)

Subtraction or addition of features feasible (through combination with other techniques, e.g., deposition, **liftoff**,…)

1.B. Photoresists

Contrast curve for ideal photoresists

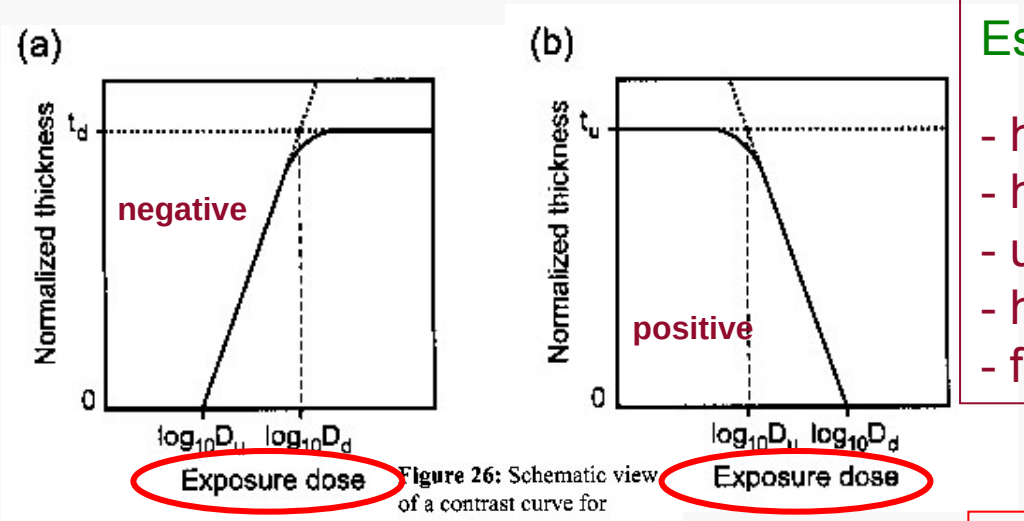


Figure 26: Schematic view of a contrast curve for (a) negative tone resist and (b) positive tone resist.

Essential requirements:

- high sticking coefficient;
- homogeneity;
- uniform thickness;
- high sensitivity to radiation (typ. UV);
- fidelity in pattern reproduction

"Sensitivity":
N = number of broken bonds = G*dose/100
With dose expressed in eV: G_{typ} ~ 1-10
(i.e., typ quantum eff ~ 1-10%)

Typical choices for photoresist:
✓Light-sensitive polymers (or organic amorphous materials)
✓UV-broken bonds modify features (protect/unprotect against etching)
✓Thickness kept below the micrometer level to improve homogeneity and reduce dose (and enhance space resolution)

A few examples of photoresists

7 Photoresist

Photoresists are also an integral part of lithography. The performance of the resist is the determining factor for the magnitude of the technology factor k_1. In general, photoresists are polymers which react when exposed to light. There are two different types of resists: With positive tone resists, the exposed areas of the resist will dissolve in the developer, with negative tone resists, the exposed areas will remain.

Positive tone resists consist of three components, a resin, which serves as a binder and establishes the mechanical properties, a photoactive compound (PAC), and a solvent to keep the resist liquid. The resin is not normally responsive to the exposure. The commonly used positive tone resist system for g- and i-line lithography is the novolac/diazonaphthoquinones (DNQ) system. The novolac is the resin material and dissolves in aqueous bases. The DNQ is the PAC, but when unexposed it acts as a dissolution inhibitor. Figure 22 shows the reaction cycle of the DNQ upon exposure. Upon exposure N_2 is split off the molecule. After a rearrangement, the molecule undergoes a reaction with the H_2O, which stems from the air. The reaction product now does not behave as a dissolution inhibitor, but as a dissolution enhancer. Therefore the exposed areas of the resist will dissolve about 100 times quicker than the unexposed areas.

Negative tone resists also consist of the three compounds: resin, photoactive compound and a solvent to keep the resist liquid. The resin consists of a cyclic synthetic rubber, which is not radiation-sensitive, but strongly soluble in the developer (non-polar organic solvents). The PAC is normally a bis-arylazide. Figure 23 shows the chemical structure of a rubber resin and a PAC. Upon exposure, the PAC dissociates into *nitrene* and N_2. These nitrene molecules are able to react with the rubber molecules, so a cross-linking between two rubber molecules can be established. Thus a three-dimesional cross-linked molecular network is formed, which is insoluble in the developer .

As device dimensions are scaled down further, the g-line steppers as well as the novolac/DNQ resists have been improved, so the features for 350 nm generation could be printed. But reaching the 250 nm generation, the illumination wavelength was shifted to 250 nm, too. However, at this wavelength novolac and DNQ do strongly absorb the light, therefore another class of resists had to be developed. Furthermore, the intensity of

Positive resist

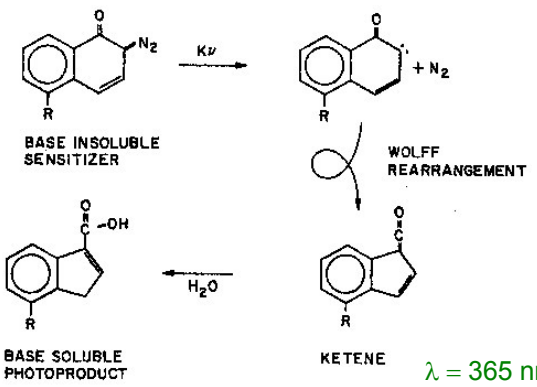


Figure 22: Exposure process of positive tone DNQ-resists [38], [39].

λ = 365 nm (Hg-lamps)

Negative resist

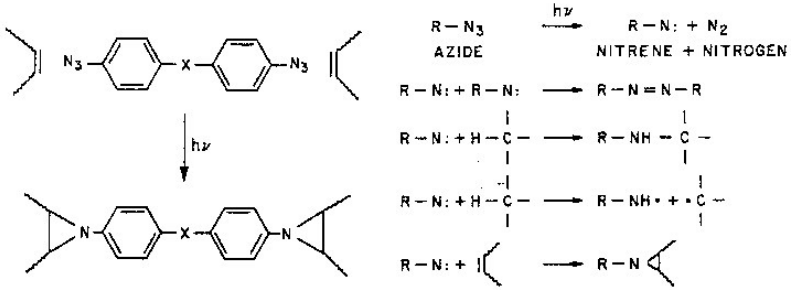


Figure 23: Reaction cycle of a negative tone resist during exposure [38], [39].

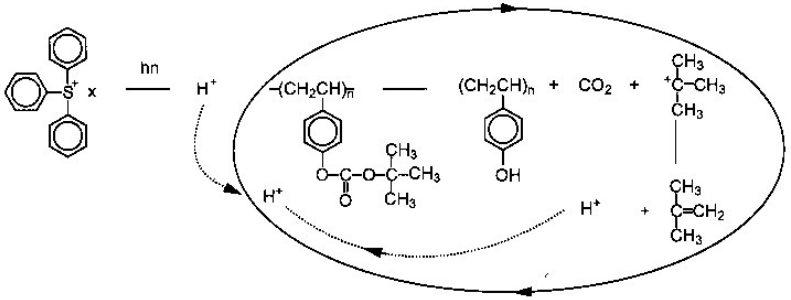
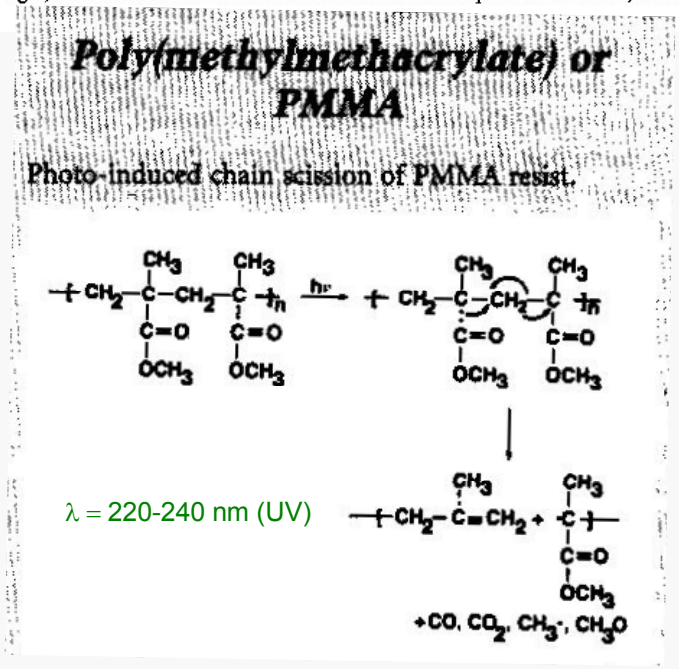


Figure 24: Chemical amplification cycle in a CAR [40].

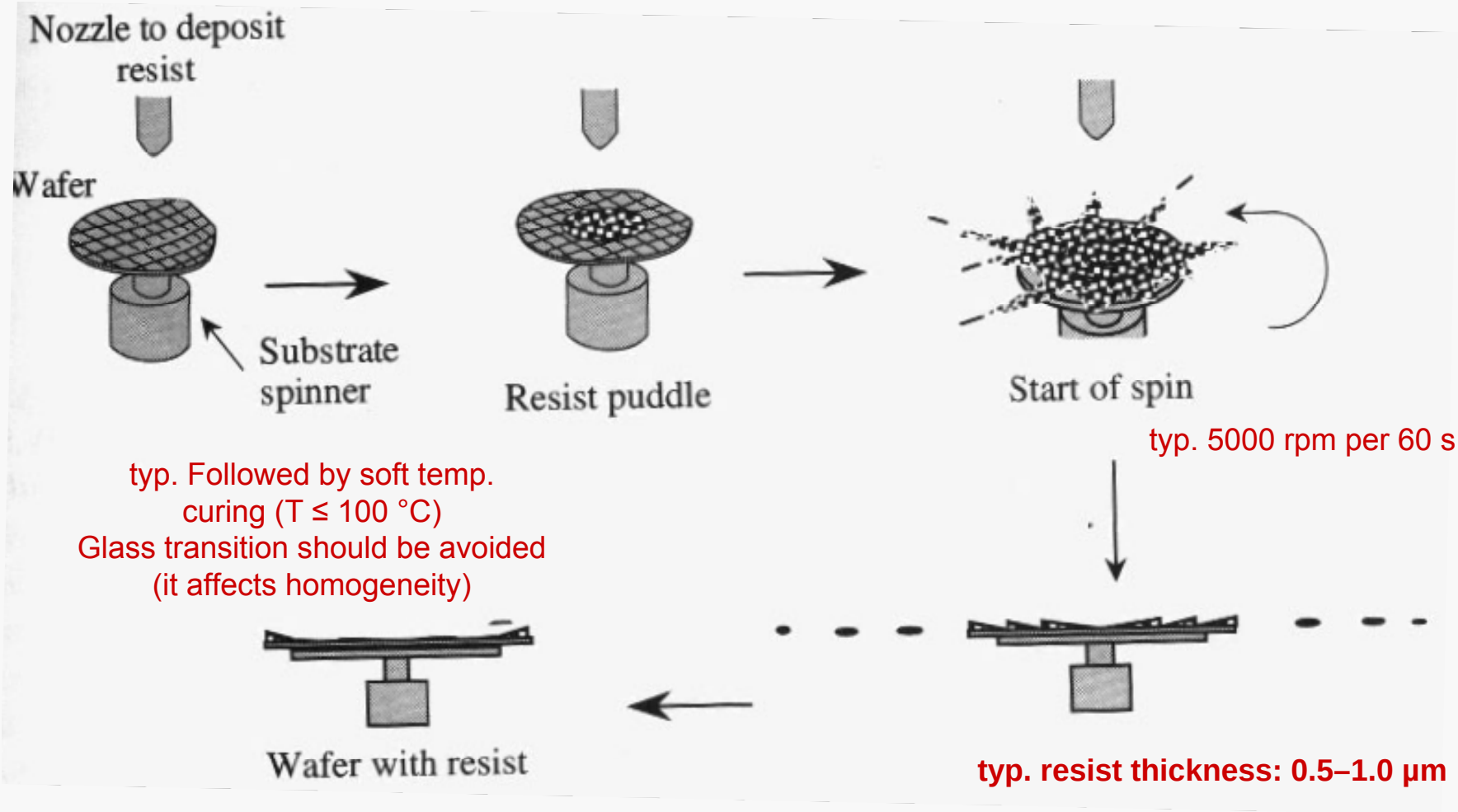
Da R. Waser Ed., Nanoelectronics and information technology (Wiley-VCH, 2003)



Well established photochemistry features

Deposition of photoresists

Most frequently used system: spin-coating (simple, scalable, effective, cheap,…)

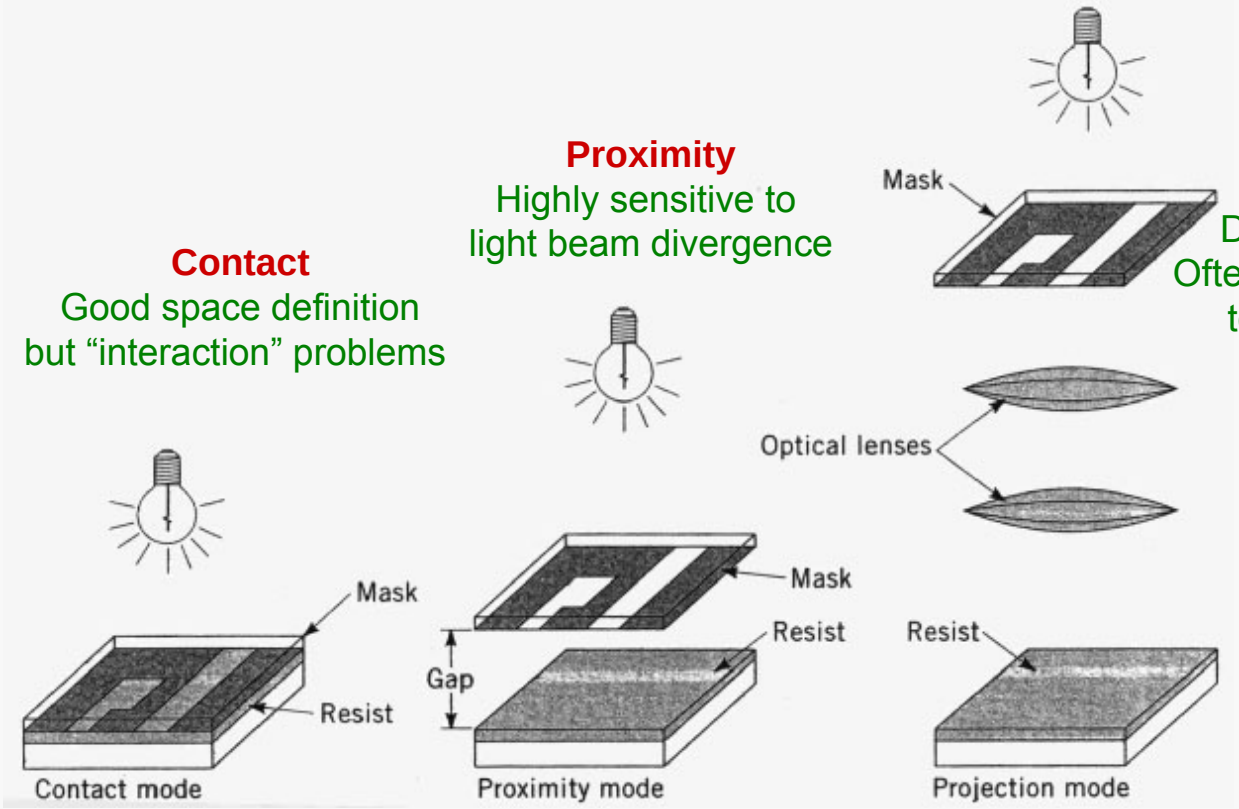


Mask configurations

Contact
Good space definition
but “interaction” problems

Proximity
Highly sensitive to
light beam divergence

Projection
High flexibility
Depends on optical quality
Often used with step-and-repeat
to cover large size wafers

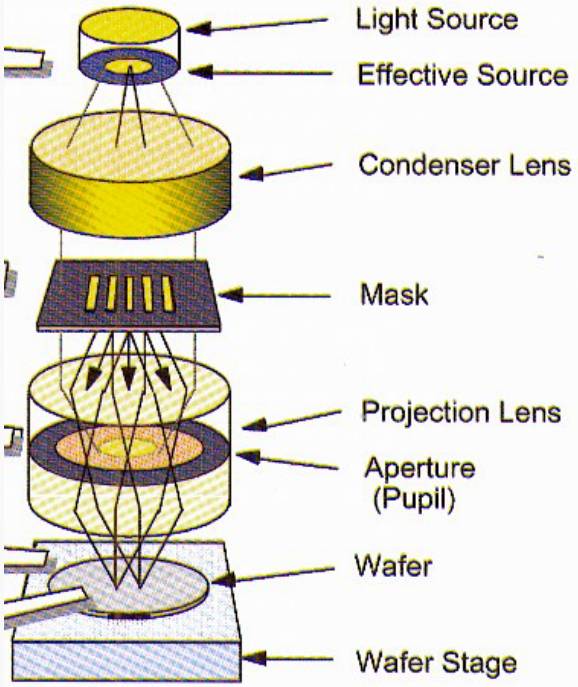


Presently: **projection** modes are most common in industrial implementations, typically implemented with “step and repeat” techniques (a small region of the substrate is impressed and then the substrate is moved to repeat the patterning over a larger area)

2. Complementarity lithography/microscopy

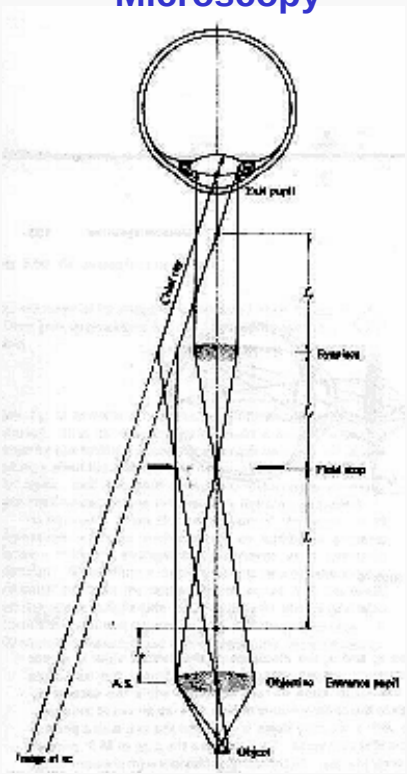
A system for optical projection lithography is clearly analogous to an optical microscope: location of light source and object are just reversed (indeed, this applies also to other lithography methods!)

Lithography



(mask aligner)

Microscopy



- **Each microscopy method corresponds to a lithography approach**
- **Issues (as, e.g., spatial resolution) are common to the two topics**

Reminders of optical microscopy

Da Hecht Zajac
Optics
Addison-Wesley (1974)

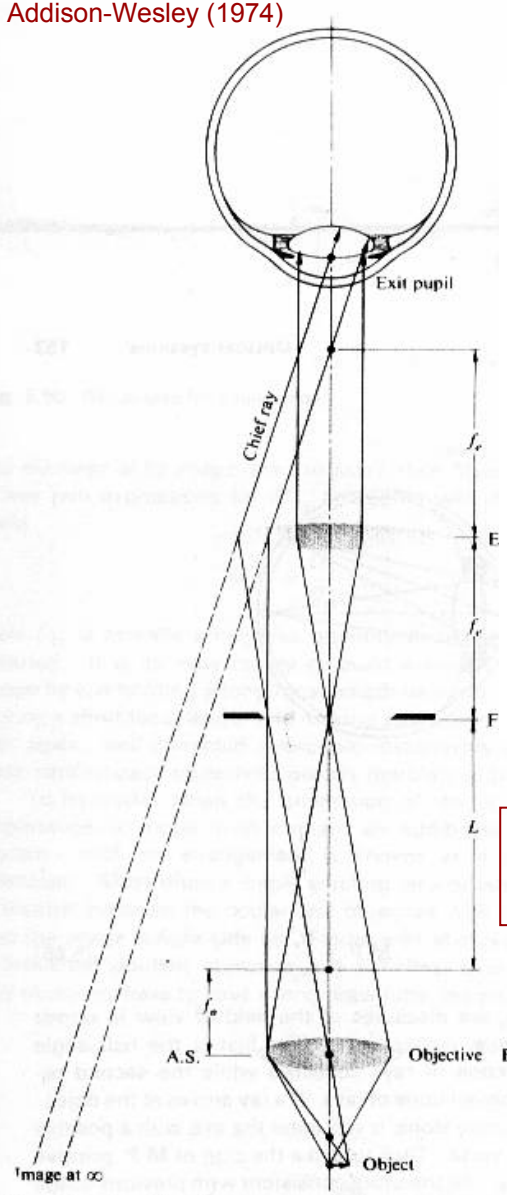


Fig. 5.86 A rudimentary compound microscope.

the magnifying power of the entire system is the product of the transverse linear magnification of the objective, M_{T_o}, and the angular magnification of the eyepiece, M_{A_e}, that is

$$\text{M.P.} = M_{T_o} M_{A_e} \qquad (5.70)$$

Recall that $M_T = -x_i/f_o$ (5.26), and with this in mind most, but not all, manufacturers design their microscopes such that the distance (corresponding to x_i) from the second focus of the objective to the first focus of the eyepiece is standardized at 160 mm. This distance, known as the *tube length*, is denoted by L in the figure. (Some authors define tube length as the image distance of the objective.) Hence, with the final image at infinity and the standard near point taken as 10 inches or 254 mm

$$\text{M.P.} = \left(-\frac{160}{f_o}\right)\left(\frac{254}{f_e}\right) \qquad (5.71)$$

and the image is inverted (M.P. < 0). Accordingly, the barrel of an objective with a focal length f_o of say 32 mm will be engraved with the markings 5× (or ×5) indicating a power of 5. Combined with a 10× eyepiece (f_e = 1 inch) the microscope M.P. would then be 50×.

Arbitrary magnification *seems* feasible (by choosing arbitrarily short focal lengths for the optical components)

Depth of field decreases as numerical aperture increases (related to *resolution*)

Magnification and depth of field depend on lens features (focal length, numerical aperture)

Da Brandon Kaplan
Microstruct. Charact.
of Materials
Wiley (1999)

Depth of field

3 Depth of Field and Depth of Focus

ce the resolution available for an object in focus in the image plane is limited by numerical aperture of the objective lens, it follows that the object need not be at exact object distance from the lens u, but may be displaced from this plane ıout sacrificing any resolution (Fig. 3.12). The distance over which the object ains in focus is defined as the *depth of field*,

$$d = \delta \tan\alpha \qquad (3.2)$$

ıre 3.12 Since the resolution is finite, the object need not be in the exact object-plane tion in order to remain in focus, and there is an allowed depth of field d. Similarly, the ıe may be observed without loss of resolution if the image plane is slightly displaced, so there is an allowed depth of focus D

where α is half the angle subtended by the objective aperture at the focal point. Similarly, the image will remain in focus if it is displaced from its geometrically defined position at a distance v from the lens. The distance over which the image remains in focus is termed the *depth of focus*, as follows:

$$D = M^2 d \qquad (3.3)$$

where M is the magnification. (Both of these expressions (equations (3.2) and (3.3)) are approximate and assume that the objective can be treated as a 'thin lens', which is never the case in a commercial instrument.) Since the resolution is given by $\delta = 0.61\lambda/\mu \sin\alpha = 0.61\lambda/NA$, it follows that the depth of field decreases as the numerical aperture increases. For the highest image resolution, the specimen should be positioned to an accuracy of better than 0.5 μm, which determines the required mechanical stability of the specimen stage.

The depth of focus is considerably less critical. Bearing in mind that a magnification of the order of 100 is necessary if all of the resolved detail is to be recorded, displacements of the order a millimetre are acceptable.

In lithography:
High magniication → short focal lenghts → increased *NA* → decreased *d* → thin resist needed!

Da Hecht Zajac, Optics
Addison-Wesley (1974)

Reminders on interference and diffraction

10.1.3 Several Coherent Oscillators

As a simple yet logical bridge between the studies of interference and diffraction, consider the arrangement of Fig. 10.6. The illustration depicts a linear array of N coherent point oscillators (or radiating antennas), which are each identical even to their polarization. For the moment, consider the oscillators to have no intrinsic phase difference, i.e. they each have the same epoch angle. The rays shown are all almost parallel, meeting at some very distant point P. If the spatial extent of the array is comparatively small, the separate wave amplitudes arriving at P will be essentially equal, having traveled nearly equal distances, that is

$$E_0(r_1) = E_0(r_2) = \cdots = E_0(r_N) = E_0(r).$$

The sum of the interfering spherical wavelets yields an electric field at P, given by the real part of

$$E = E_0(r)e^{i(kr_1 - \omega t)} + E_0(r)e^{i(kr_2 - \omega t)} + \cdots + E_0(r)e^{i(kr_N - \omega t)}. \tag{10.1}$$

It should be clear, from Section 9.1, that we need not be concerned with the vector nature of the electric field for this configuration. Now then

$$E = E_0(r)e^{-i\omega t}e^{ikr_1}[1 + e^{ik(r_2 - r_1)} + e^{ik(r_3 - r_1)} + \cdots + e^{ik(r_N - r_1)}].$$

The phase difference between adjacent sources is obtained from the expression $\delta = k_0\Lambda$ and since $\Lambda = nd\sin\theta$, in a medium of index n, $\delta = kd\sin\theta$. Making use of Fig. 10.6, it follows that $\delta = k(r_2 - r_1)$, $2\delta = k(r_3 - r_1)$ etc. Thus the field at P may be written as

$$E = E_0(r)e^{-i\omega t}e^{ikr_1}[1 + (e^{i\delta}) + (e^{i\delta})^2 + (e^{i\delta})^3 + \cdots (e^{i\delta})^{N-1}]. \tag{10.2}$$

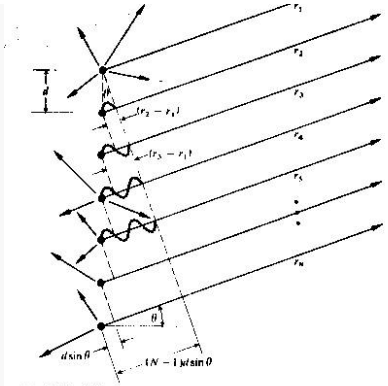


Fig. 10.6 A linear array of in-phase coherent oscillators. Note that at the angle shown $\delta = \pi$ while at $\theta = 0$ δ would be zero.

The bracketed geometric series has the value

$$(e^{i\delta N} - 1)/(e^{i\delta} - 1)$$

which can be rearranged into the form

$$\frac{e^{iN\delta/2}[e^{iN\delta/2} - e^{-iN\delta/2}]}{e^{i\delta/2}[e^{i\delta/2} - e^{-i\delta/2}]}$$

or equivalently

$$e^{i(N-1)\delta/2}\left[\frac{\sin N\delta/2}{\sin\delta/2}\right].$$

The field then becomes

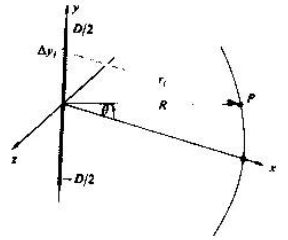


Fig. 10.8 A coherent line source.

$$E = E_0(r)e^{-i\omega t}e^{i[kr_1 + (N-1)\delta/2]}\left(\frac{\sin N\delta/2}{\sin\delta/2}\right). \tag{10.3}$$

Notice that if we define R to be the distance from the center of the line of oscillators to the point P, that is

$$R = \tfrac{1}{2}(N-1)d\sin\theta + r_1,$$

then Eq. (10.3) takes on the form

$$E = E_0(r)e^{i(kR - \omega t)}\left(\frac{\sin N\delta/2}{\sin\delta/2}\right). \tag{10.4}$$

Finally, then, the flux-density distribution within the diffraction pattern due to N coherent, identical, distant point sources in a linear array is proportional to $EE^*/2$ for complex E or

$$I = I_0\frac{\sin^2(N\delta/2)}{\sin^2(\delta/2)}, \tag{10.5}$$

where I_0 is the flux density from any single source arriving at P (see Problem 10.2 for a graphical derivation of the irradiance). For $N = 0$, $I = 0$, for $N = 1$, $I = I_0$, and for $N = 2$, $I = 4I_0\cos^2(\delta/2)$ in accord with Eq. (9.6). The functional dependence of I on θ is more apparent in the form

$$I = I_0\frac{\sin^2[N(kd/2)\sin\theta]}{\sin^2[(kd/2)\sin\theta]}. \tag{10.6}$$

The $\sin^2[N(kd/2)\sin\theta]$ term undergoes rapid fluctuations, while the function modulating it, $\{\sin[(kd/2)\sin\theta]\}^{-2}$, varies relatively slowly. The combined expression gives rise to a series of sharp principal peaks separated by small subsidiary maxima. The principal maxima occur in directions θ_m such that $\delta = 2m\pi$ where $m = 0, \pm1, \pm2, \ldots$. Because $\delta = kd\sin\theta$

$$d\sin\theta_m = m\lambda. \tag{10.7}$$

Since $[\sin^2 N\delta/2]/[\sin^2\delta/2] = N^2$ for $\delta = 2m\pi$ (from

Fraunhofer diffraction

L'Hôspital's rule) the principal maxima have values $N^2 I_0$. This is to be expected inasmuch as all of the oscillators are in phase at that orientation. The system will radiate a maximum in a direction perpendicular to the array ($m = 0$, $\theta_0 = 0$ and π). As θ increases, δ increases and I falls off to zero at $N\delta/2 = \pi$, its first minimum. Note that if $d < \lambda$ in Eq. (10.7), only the $m = 0$ or zero-order principal maximum exists. *If we were looking at an idealized line source of electron-oscillators separated by atomic distances, we could expect only that one principal maximum in the light field.*

The antenna array of Fig. 10.7 can then transmit radiation in the narrow beam or lobe corresponding to a principal maximum (the parabolic dishes shown reflect into the forward direction and the radiation pattern is no longer symmetrical around the common axis.) Suppose that we have a system in which we can introduce an intrinsic phase shift of ε between adjacent oscillators. In that case

$$\delta = kd \sin\theta + \varepsilon;$$

the various principal maxima will occur at new angles

$$d \sin\theta_m = m\lambda - \varepsilon/k.$$

Concentrating on the central maximum $m = 0$, its orientation θ_0 can be varied at will by merely adjusting the value of ε.

The principle of reversibility, which states that without absorption, wave motion is reversible, leads to the same field pattern for an antenna used as either a transmitter or receiver. The array, functioning as a radio telescope, can therefore be "pointed" by combining the output from the individual antennas with an appropriate phase shift, ε, introduced between each of them. For a given ε the output of the system corresponds to the signal impinging on the array from a specific direction in space.

Figure 10.7 is a photograph of the first multiple radio interferometer designed by W. N. Christiansen and built in Australia in 1951. It consists of 32 parabolic antennas, each 2 m in diameter, designed to function in phase at the wavelength of the 21 cm hydrogen emission line. The antennas are arranged along an east–west baseline with 7 m separating each one. This particular array utilizes the earth's rotation as the scanning mechanism.

Examine Fig. 10.8 which depicts an idealized line source of electron-oscillators (e.g., the secondary sources of the Huygens–Fresnel principle for a long slit whose width is much less than λ illuminated by plane waves). Each point emits a spherical wavelet which we write as

$$E = \left(\frac{\mathcal{E}_0}{r}\right) \sin(\omega t - kr)$$

explicitly indicating the inverse r-dependence of the amplitude. The quantity $\mathcal{E}_0$ is said to be the *source strength*. The present situation is distinct from that of Fig. 10.6 in that now the sources are very weak, their number, N, is tremendously large and the separation between them vanishingly small. A minute, but finite segment of the array Δy_i, will contain $\Delta y_i(N/D)$ sources where D is the entire length of the array. Imagine then that the array is divided up into M such segments, i.e., i goes from 1 to M. The contribution to the electric field intensity at P from the ith segment is accordingly

$$E_i = \left(\frac{\mathcal{E}_0}{r_i}\right) \sin(\omega t - kr_i) \left(\frac{N\Delta y_i}{D}\right)$$

provided that Δy_i is so small that the oscillators within it have a negligible relative phase difference (r_i = constant) and their fields simply add constructively. We can cause the array to become a continuous (coherent) line source by letting N approach infinity. This description, besides being fairly realistic on a macroscopic scale, also allows the use of the calculus for more complicated geometries. Certainly as N approaches infinity, the source strengths of the individual oscillators must diminish to near zero if the total output is to be finite. We can therefore define a constant $\mathcal{E}_L$ as the *source strength per unit length* of the array, that is

$$\mathcal{E}_L \equiv \frac{1}{D} \lim_{N \to \infty} (\mathcal{E}_0 N). \tag{10.8}$$

The net field at P from all M segments is

$$E = \sum_{i=1}^{M} \frac{\mathcal{E}_L}{r_i} \sin(\omega t - kr_i)\, \Delta y_i.$$

For a continuous line source Δy_i can become infinitesimal ($M \to \infty$) and the summation is then transformed into a definite integral

$$E = \mathcal{E}_L \int_{-D/2}^{+D/2} \frac{\sin(\omega t - kr)}{r}\, dy, \tag{10.9}$$

where $r = r(y)$. The approximations used to evaluate Eq. (10.9) must depend on the position of P with respect to the array and will therefore make the distinction between Fraunhofer and Fresnel diffraction. The coherent *optical* line source does not now exist as a physical entity but we will make good use of it as a mathematical device.

10.2 FRAUNHOFER DIFFRACTION

10.2.1 The Single Slit

Return to Fig. 10.8 where now the point of observation is very distant from the coherent line source and $R \gg D$. Under these circumstances $r(y)$ never deviates appreciably from its midpoint value R so that the quantity $(\mathcal{E}_L/R)$ at P is essentially constant for all elements dy. It follows from Eq. (10.9) that the field at P due to the differential segment of the source dy is

$$dE = \frac{\mathcal{E}_L}{R} \sin(\omega t - kr)\, dy, \tag{10.10}$$

where $(\mathcal{E}_L/R)\, dy$ is the amplitude of the wave. Notice that the phase is very much more sensitive to variations in $r(y)$ than is the amplitude so that we will have to be more careful about introducing approximations into it. We can expand $r(y)$, in precisely the same manner as was done in Problem (9.4), to get it as an explicit function of y, thus

$$r = R - y \sin\theta + (y^2/2R) \cos^2\theta + \cdots, \tag{10.11}$$

where θ is measured from the xz-plane. The third term can be ignored so long as its contribution to the phase is insignificant even when $y = \pm D/2$, i.e. $(\pi D^2/4\lambda R) \cos^2\theta$ must be negligible. This will be true for all values of θ when R is adequately large and we again have the Fraunhofer condition. The distance r is then linear in y. Substituting into Eq. (10.10) and integrating leads to

$$E = \frac{\mathcal{E}_L}{R} \int_{-D/2}^{+D/2} \sin[\omega t - k(R - y\sin\theta)]\, dy, \tag{10.12}$$

and finally

$$E = \frac{\mathcal{E}_L D}{R} \frac{\sin[(kD/2)\sin\theta]}{(kD/2)\sin\theta} \sin(\omega t - kR). \tag{10.13}$$

To simplify the appearance of things let

$$\beta \equiv (kD/2)\sin\theta \tag{10.14}$$

so that

$$E = \frac{\mathcal{E}_L D}{R} \left(\frac{\sin\beta}{\beta}\right) \sin(\omega t - kR). \tag{10.15}$$

The quantity most readily measured is the irradiance (forgetting the constants) $I(\theta) = \langle E^2 \rangle$ or

$$I(\theta) = \frac{1}{2}\left(\frac{\mathcal{E}_L D}{R}\right)^2 \left(\frac{\sin\beta}{\beta}\right)^2, \tag{10.16}$$

where $\langle \sin^2(\omega t - kR) \rangle = \frac{1}{2}$. When $\theta = 0$, $\sin\beta/\beta = 1$ and $I(\theta) = I(0)$ which corresponds to the *principal maximum*. *The irradiance resulting from an idealized coherent line source in the Fraunhofer approximation* is then

$$I(\theta) = I(0) \left(\frac{\sin\beta}{\beta}\right)^2, \tag{10.17}$$

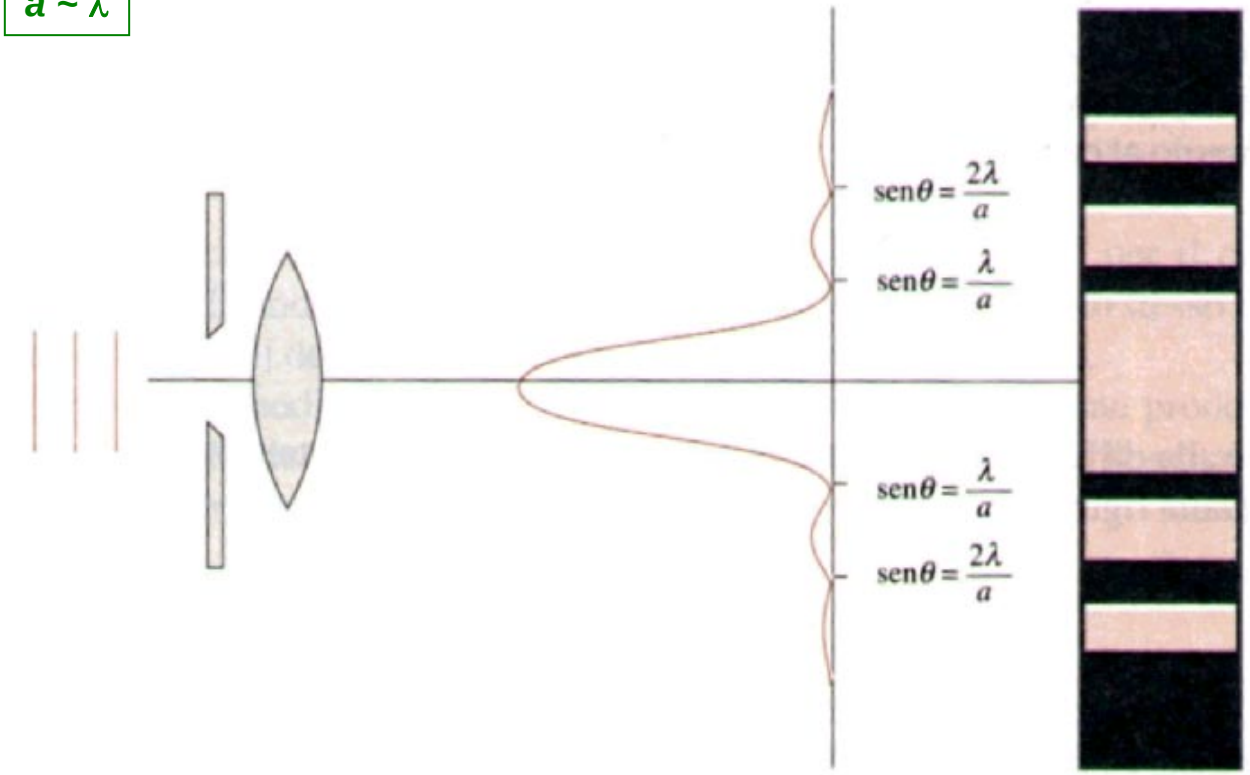
or using the *sinc function* (Section 7.9, and Table 1 of the Appendix)

$$I(\theta) = I(0)\, \mathrm{sinc}^2\beta.$$

There is symmetry about the y-axis and this expression holds for θ measured in any plane containing that axis.

Effects of diffraction

$a \sim \lambda$



"Diffusion cone":
$\text{Sin}\theta \sim \lambda/a$

Diffraction ripples:
$I \sim I_0 (\sin\alpha/\alpha)^2$
with:
$\alpha = \pi\, a\, \sin\theta\, /\, \lambda$

Optical diffraction is for sure a *fundamental* limiting factor in optical microscopy and lithography

Criteria for space resolution (in optical microscopy)

3.1.2.1 POINT-SOURCE ABBE IMAGE

The calculated intensity distribution assumes a parallel beam of light travelling along the axis of a thin lens and brought to a focus at the focal distance (Fig. 3.8). For the *cylindrically symmetric* case, the ratio of the peak intensities for the primary and secondary peaks in the intensity distribution is *ca* 9 : 1, while the width of the primary peak is given by the *Abbe equation* as follows:

Abbe

$$\delta = 0.61 \frac{\lambda}{\mu \sin \alpha} \qquad (3.1)$$

where λ is the wavelength of the radiation, α is the aperture (half-angle) of the lens (determined by the ratio of the lens radius to its focal length), and μ is the refractive index of the medium between the lens and the focal point ($\mu \approx 1$ for air).

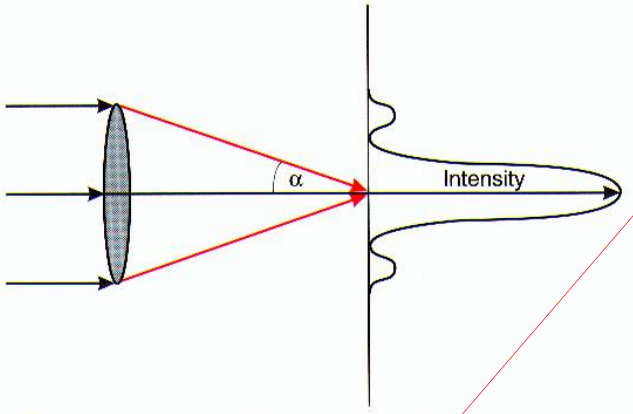
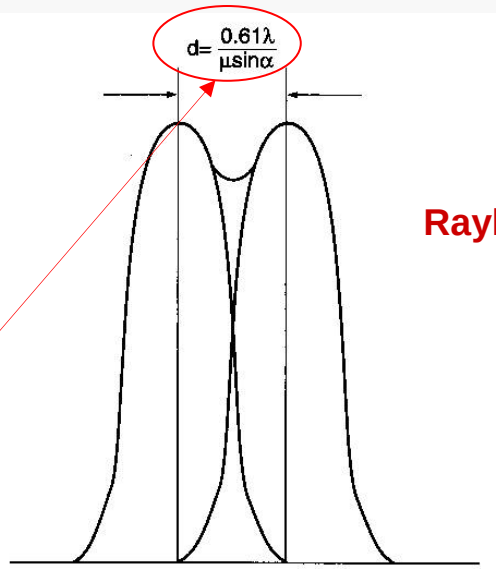


Figure 3.8 The Abbe equation gives the width of the first intensity peak for the image of point object at infinity in terms of the angular aperture of the lens α and the wavelength of tl radiation λ

Da Brandon Kaplan
Microstruct. Charact.
of Materials
Wiley (1999)



Rayleigh

Figure 3.10 The Raleigh resolution criterion requires that two point sources at infinity have an angular separation which is sufficient to place the maximum intensity of the primary image peak of one source at the position of the first minimum of the second

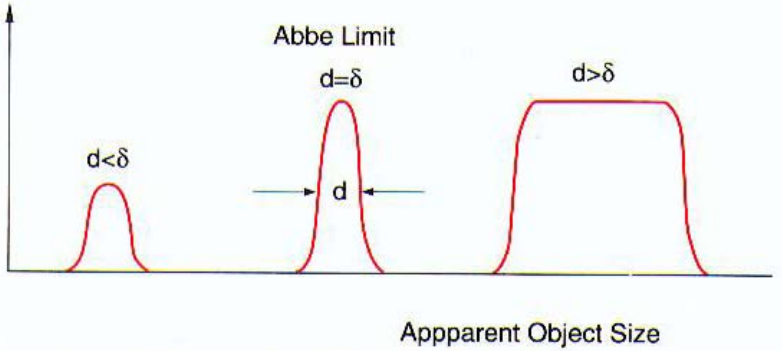


Figure 3.11 Large objects of diameter d are blurred by the diffraction limit δ derived from the Abbe relationship, but objects smaller than the Abbe width are still detectable in the microscope, although the intensity is reduced and they have an apparent width given by the Abbe equation

Maximum achievable space resolution $d \sim 0.61\,\lambda / (NA) < \lambda/2$
(NA: numerical aperture of the optical system: NA = $n \sin\alpha$, with n refractive index)

A (qualitative) look at Fourier-transform optics

A surface pattern $f(x,y)$ (i.e., an object to be imaged!) can be always described in terms of a (2D) Fourier superposition

Fourier components are function of the reciprocal space k_X k_Y :

$$f(x,y) = \int f'(k_X, k_Y) \exp(-ik_X x) \exp(-ik_Y y)\, dk_X\, dk_Y$$

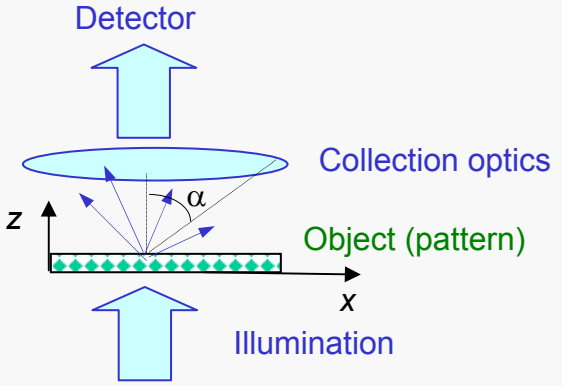
Large k_X (and k_Y) values imply large angular displacement

⇩

A faithful pattern reproduction (i.e., an accurate image) implies collecting the largest number of k_X and k_Y

⇩

Large numerical aperture ($NA=nSin\theta$) needed for high resolution



but

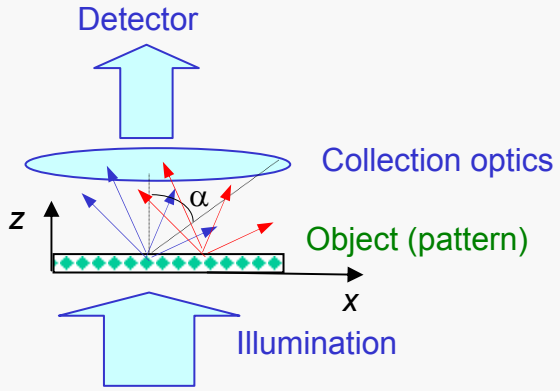
In non point-like illumination schemes (conventional) rays emitted from different points of the surface can be collected

⇩

Large numerical aperture leads to sensitivity to “stray light”

⇩

Contrast falls down and high space resolution is hampered



Contrast (i.e., signal-to-noise) affects the ultimate resolution

3. Improving resolution: optical confocal microscopy I

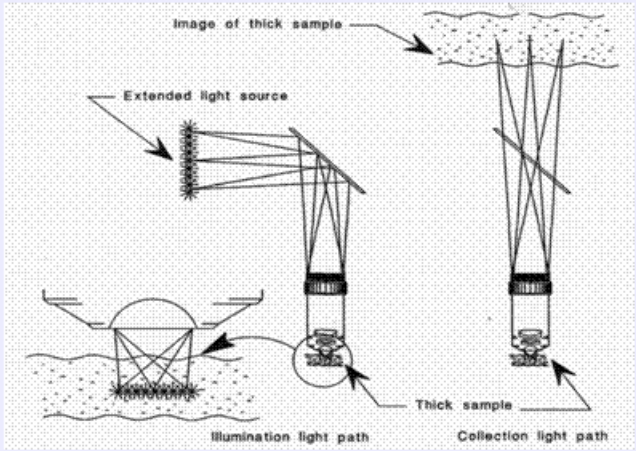
In practical terms, space resolution is also a matter of *contrast*

In all microscopies, methods exist to improve the contrast, so enhancing the resolution

For instance, in optical microscopy **confocal** systems have been developed with a space resolution on the order of 200 nm (for the visible light), i.e., close to the diffraction limit

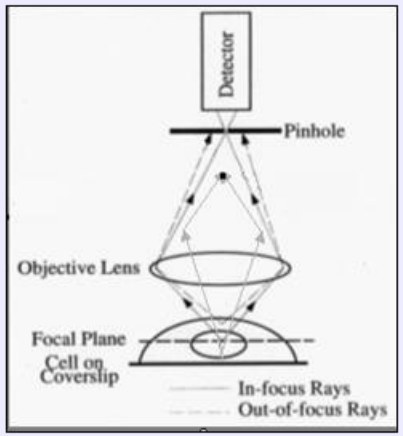
“Stray light” effects are removed and the contrast is enhanced

Extended Light Versus Point Light Source Illumination



In conventional wide field microscope, ordinary extended light is used as light source, the specimen is lit laterally and vertically at the same time as shown in the illustration. The resulting image is affected by all the lit spots from the whole illuminated field, although it is centered at a given focal plane and local spot. These illuminated dots interfere with each other laterally and the stray light compromise image contrast. Image contrast, defined as the difference between the minimum and maximum intensity of two points in the image, is an important factor for an optical device to achieve its resolution, without proper contrast, the signal has little difference with background and the resolution of the an optical lens can not be realized. Improved contrast helps an optical device to reach its maximum resolution.

In another configuration, a plate with a small hole called pinhole is placed before the image detecting device like below:



In this configuration, light from under-focal-plane will be focused at a plane behind the pinhole such is blocked away by the pinhole plate. The light from above-focal-plane will be focused before the pinhole and is blocked away by the pinhole too. Only the light from focal plane is just focused at the pinhole thus can reach the image detector. This process simulates what you do with a microtome to cut some unwanted tissue away, but you do it here optically, this is so called "**optical sectioning**".

The size of pinhole determines how thick an optical slice will be. The smaller the pinhole, the thinner the slice. But the thickness will not go down indefinitely. It is also limited by all those factors affecting resolution of the lens: the wave length of light, Numerical aperture of the lens, reflecting index of media, together with pinhole size, the z-resolution is usually 2 times worse than lateral resolution of an objective. For a lens of 1.4 NA, blue light at 488 nm, the lateral resolution is 200 nm, the achievable optical section thickness is about 400 nm.

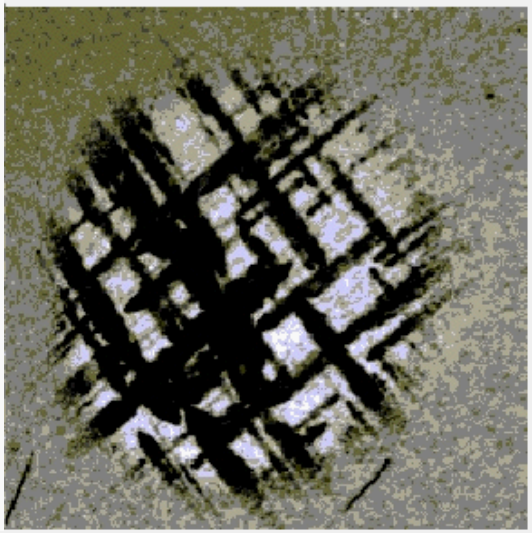
http://www.hi.helsinki.fi/amu/AMU%20Cf_tut/Cf_tut_part1-3.htm

Optical confocal microscopy II

Confocal microscopy often used for biological samples with fluorescent markers (e.g., quantum dots)

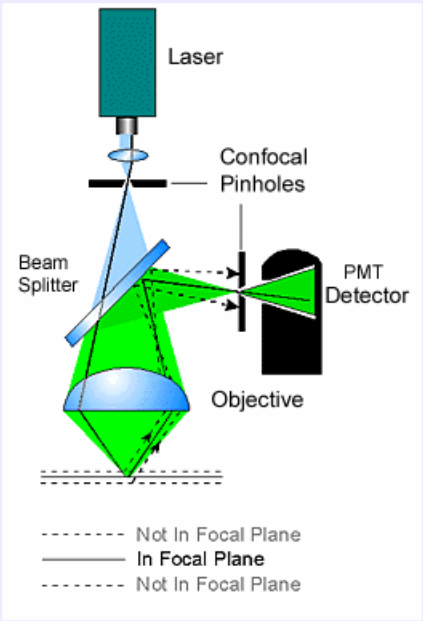
3D mapping capabilities can be added by moving the system along the focal axis

Frequently **2-photon excitation is used to further enhance space resolution** (due to the nonlinear dependence of the absorption probability on the exciting intensity)



"Segregation" of quantum wells

Confocal photoluminescence images of a II-VI laser structure sh
ZnCdSe/ZnSSe/ZnMgSSe separate confinement heterostructure:



In practical, a point-like light source is achieved by using a laser light passing through a illumination pinhole. This point-like light source is directed to the specimen by a beam splitter (or AOBS in Leica's BSP-free system) to form a point-like illumination in the specimen. The point-illumination move or scan on the specimen by the help of a scanner. The reflected emission light from specimen's focal plane passes through the detecting pinhole and form point-like image on detector PMT (photon multiply tube).

PMT converts detected photon into electron. It is possible to amplify weak signal by manipulating the voltage (gain) on the tube. It is also possible to cut off background signal by set certain threshold (Offset) on the tube.

PMT has large active area to receive photons thus high saturate point, and PMT has low dark current thus low background, together, which provide high dynamic range that is defined as the ratio of maximum allowed intensity / dark current. Besides, PMT has very high refresh rate since there is no charge accumulate on it. It detects event at nano- seconds level.

Taking together, in confocal system:

- A point light source for illumination
- A point light focus within the specimen
- A pinhole at the image detecting plane

Scanning technique!

These **three points** are **optically conjugated together** and aligned accurately to each other in the light path of image formation, this is **confocal**. Confocal effects result in supression of out-of-focal-plane light, supression of stray light in the final image

Confocal images have following features:

- void of interference from lateral stray light: higher contrast.
- void of supperimpose of out-of-focal-plane signal: less blur, sharper image.
- images derived from optically sectioned slices (depth discrimination)
- Improved resolution (theoretically) due to better wave-optical performance.

Is confocal effect a free cake?
No, confocal effect is obtained at a cost of greatly reduced detecting volume (total signal amount), increase vulnerability to noise, reduced dynamic range, etc. For detailed discussion, refer section 6: Optical sectioning.

Space resolution in optical lithography I

Confocal schemes can be used to push resolution of an optical microscope towards the fundamental (diffracion) limit, but technique is scanning (no more parallel) and requires transparent samples ⇒ **small interest in lithography**

Half-pitch, critical size, space resolution in lithography

wavelike nature of light leads to diffraction, reducing contrast between nominally bright and nominally dark areas of the pattern. In analyzing this phenomenon with the tools of classical optics, including Fourier decomposition,[2] it quickly becomes apparent that the parameter controlling the image quality is the spatial frequency, or pitch, rather than the size of the smallest feature.

The Rayleigh equation is a commonly used rule of thumb in optical lithography, relating the half-pitch (HP) to the vacuum wavelength of the exposure light (λ), and the numerical aperture of the imaging system, $NA = n\sin\theta$, where n is the refractive index of the medium above the photoresist-coated wafer and θ is the half-angle of the converging beam in the image plane (see Figure 1):

$$HP = k_1\lambda/NA. \quad (1)$$

In this equation, k_1 is a parameter determined by the illumination conditions, the structures and materials of the photomask, the quality of the imaging optics, and the performance of the photoresist.

MRS Bull., vol. 30 (dec. 2005)

Mass-produced semiconductor manufacturing entered the era of nanopatterning when the smallest feature sizes crossed the 100-nm threshold. For the last two years, the "half-pitch" (half the distance between the centers of adjacent features in periodic structures) has been nominally 90 nm, and according to the *International Roadmap for Semiconductors* (ITRS), the size reductions will continue for at least one more decade.[1] The so-called nodes act as signposts: 65 nm in 2007, 45 nm in 2010, 32 nm in 2013, and so on. To understand the true state of lithography, it is important to understand what these numbers mean. The nodes represent the nominal half-pitch of the densest features on the chip. The smallest features ("critical dimensions," or CDs) will be equal to this half-pitch only when dense patterns of equal lines and spaces are formed. Such a geometry is common in memory devices such as dynamic random-access memories (DRAMs). In microprocessors, however, the patterns are more typically semi-isolated lines, where the width of the lines is much smaller than the width of the spaces. Thus, the CDs can be much smaller than the half-pitch. For example, at the 65-nm node, the CD in a microprocessor may be only 45 nm, with each line separated by 85-nm spaces, for a total pitch of 130 nm. Adding further complication, the definition of a node by a particular semiconductor manufacturer may vary from that of the ITRS, depending on the target product. An accurate comparison of lithographies across the semiconductor industry must thus take into account these sometimes hidden assumptions.

This seeming hair-splitting between half-pitch and CD is dictated by the fundamental constraints and challenges facing projection optical lithography. By far the most common patterning technology, optical lithography clearly has advantages over its competitors, particularly in high throughput, allowing economical mass production and defect control and enabling acceptable product yield. Its

Da M. Madou,
Fundamentals of microfab.,
CRC (1997)

Space resolution in optical lithography II

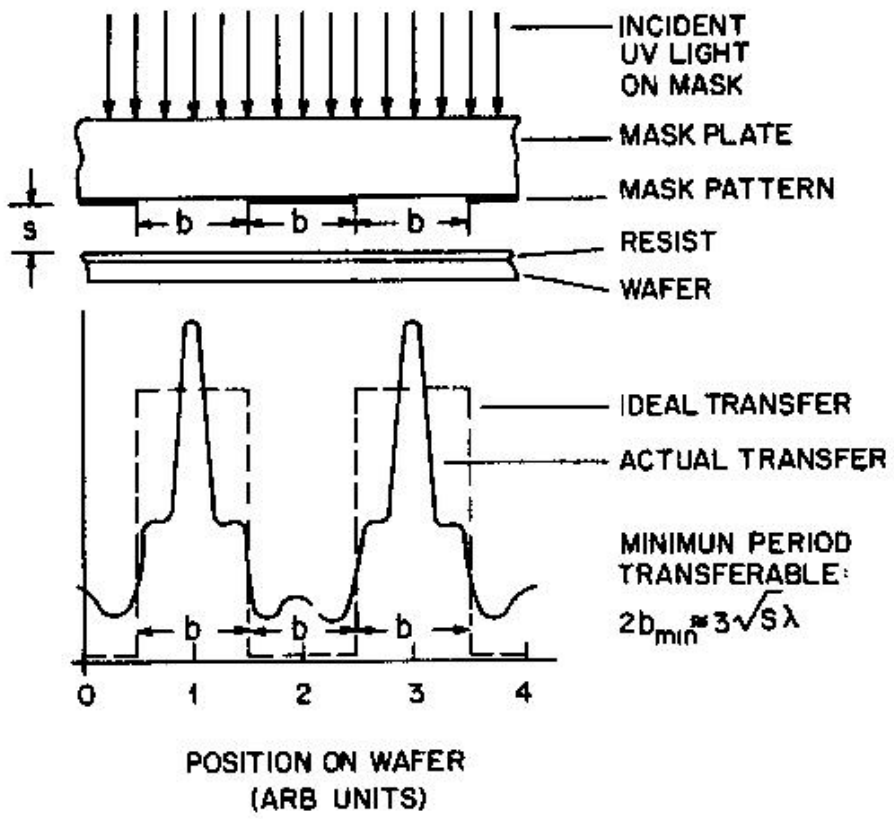


Figure 1.7 Light distribution profiles on a photoresist surface after light passed through a mask containing an equal line and space grating. (From Willson, C. G., in *Introduction to Microlithography*, Thompson, L. F., Willson, C. G., and Bowden, M. J., Eds., American Chemical Society, Washington, D.C., 1994. With permission.)

Further *technological* limitations:

Mask/substrate distance (divergence);
Resist thickness and depth of field;
Resist homogeneity

Empirical formula:

$$2b_{min} = 3\sqrt{\lambda\left(s + \frac{Z}{2}\right)} \qquad 1.12$$

where b_{min} stands for half the grating period, s for the gap between the mask and the photoresist surface, λ for the wavelength of the exposing radiation, and z for the photoresist thickness.

Example:
λ = 350 nm, s = 5 μm, Z = 0.5 μm
$\Rightarrow b_{min} > 2$ μm!!!

Care must be put even to approach the diffraction limit

More on space resolution

2 Optical Lithography

Optical lithography is the most important type of lithography. Originally the name referred to lithography using light with wavelength in the visible range. Nevertheless, gradually, the wavelength was driven down to 193 nm, which is used in semiconductor production nowadays, and even shorter wavelengths down to the sub-nm rage are under investigation.

The key issue of lithography is the resolution of the system, and hence the size of the smallest feature (minimum feature size: *MFS*) which can be defined on the sample. This *MFS* depends on the illumination method, the illumination wavelength λ, on the materials of the optical system and the resist used. In Sec. 2.1 the different illumination methods and their physical resolution limits are addressed, in Sec. 2.2 the wavelengths and the light sources are discussed, also for wavelengths below 15 nm, while lithography with these wavelengths is discussed in Sec. 3 and 4, and in Sec. 2.3 the materials and the forms of the optical system are dealt with.

2.1 Illumination Methods and Resolution Limits

Figure 3 shows a schematic view of the three different illumination methods *contact, proximity* and *projection lithography*. With all three, the light emitted by a light source passes a condenser optics so that a parallel beam is formed. With contact lithography, mask and sample are pressed together so that the mask is in close contact to the resist (Figure 3a). The resolution is limited by deflection and is expressed by the *MFS* which can be obtained. For contact lithography this is: $MFS = \sqrt{d \cdot \lambda}$, where d is the resist thickness and λ the wavelength. For a resist thickness of 1 µm and a wavelength of about 400 nm, this yields a minimum feature size of 600 nm. The major drawback of this method is that the quality of the mask suffers from contact to the resist, leading to failures in the structure. To avoid this problem, the second method was developed (Figure 3b). With *proximity lithography* there is a defined proximity gap g between sample and mask, so there is no deterioration of the mask. The drawback is the poorer resolution limit, which is proportional to $\sqrt{(d+g)\cdot\lambda}$. With same figures as above and a proximity gap of 10 µm, the *MFS* is 2 µm.

The method used today in industrial production is so-called *projection lithography* (Figure 3c). Here not the shadow of the mask is transferred to the sample as with the two other methods, but a picture of the mask is projected onto the sample. Therefore after passing the mask, the light is bundled by an optical system. The mask is not in contact with the sample, so there is no deterioration as in contact lithography, but the resolution is better than in proximity lithography. Furthermore it is possible to reduce the picture so the patterns on the mask are allowed to be bigger than the patterns on the sample. This is advantageous for mask fabrication: Errors are also reduced. If it is possible to obtain masks with an accuracy of 100 nm, then the error for a structure of 500 nm to be transferred onto a sample is 20 %, if it is transferred one by one. If the picture is reduced 4 times, then for a 500 nm feature on the sample, the feature on the mask has to be 2 µm; therefore the mask error is only 5 %. Because of the reduction, the wafer is not exposed in one exposure, but in several . This is done by so-called steppers, in which the wafer is adjusted under the mask by an x-y-table. The stepper moves the wafer from one exposure position to the next, while the mask is not moved.

In projection lithography the limiting factor to the MFS is diffraction. Consider a slit width b which is illuminated by a monochromatic plane wave. What will the intensity distribution look like on a screen at a distance l behind the slit? Therefore consider two Huygens waves, one from the lower rim of the slit, one from the middle. There will be an optical path difference between these two Huygens waves, depending on the angle of propagation Θ. The magnitude of the path difference (*PD*) is:

$$PD = \frac{b}{2}\sin(\Theta) \tag{1}$$

The two Huygens waves will interfere destructively if the *PD* is an odd multiple of the half wavelength:

$$\frac{b}{2}\sin(\Theta_{min}) = (2m+1)\cdot\frac{\lambda}{2} \qquad \text{with } m = 0, \pm 1, \pm 2, \ldots \tag{2}$$

Under this condition, the Huygens waves from the lower part of the slit will interfere destructively with the ones from the upper part. At the angle Θ_{min} there is a minimum of intensity.

The Huygens waves do interfere constructively resulting in a maximum of intensity when:

$$\frac{b}{2}\sin(\Theta_{max}) = m\lambda \qquad \text{with } m = 0, \pm 1, \pm 2, \ldots \text{ holds.} \tag{3}$$

In lithography the diffraction patterns of several structures are superimposed so the question leading to the MFS is the question of when two structures can be resolved. The first approach is given by the Rayleigh criterion [3]. When light coming from a point source passes an optical system a blurred diffraction pattern – the Airy disc – occurs. The Rayleigh criterion says that two ideal point sources (e.g. stars) can be resolved when the intensity maximum of the one Airy disc is in the first minimum of the other, so *MFS* is given as:

$$MFS = 0.61\cdot\frac{\lambda}{NA} \tag{4}$$

where *NA* is the numerical aperture of the optical system. Nevertheless the Rayleigh criterion is just a first approach to the *MFS* in microlithography. The mask patterns are not independent (i.e. incoherent) ideal point sources, on the contrary they have a finite width and the light is partially coherent. Nevertheless, the form of the criterion gives the right dependences. If the wavelength is decreased by 10 % or the *NA* is increased by 10 %, the *MFS* is improved by 10 %. Furthermore, it was derived only by properties of the optics although the photoresist also affects the MFS. Therefore more generally, the criterion is written as:

$$MFS = k_1\cdot\frac{\lambda}{NA} \tag{5}$$

where k_1 is a constant (typically 0.5 – 0.9), which accounts for non-ideal behaviour of the equipment (e.g. lens errors) and the influences which do not come from the optics (resist, resist processing, shape of the imaged structures,...). Therefore k_1 is called the technology constant.

As a comparison, for a technology constant of 0.7 and a numerical aperture of 0.7, which are commonly used figures, the *MFS* is in the order of the wavelength λ. So it is better by about a factor of 0.66 than the *MFS* of contact printing.

Modulation transfer function

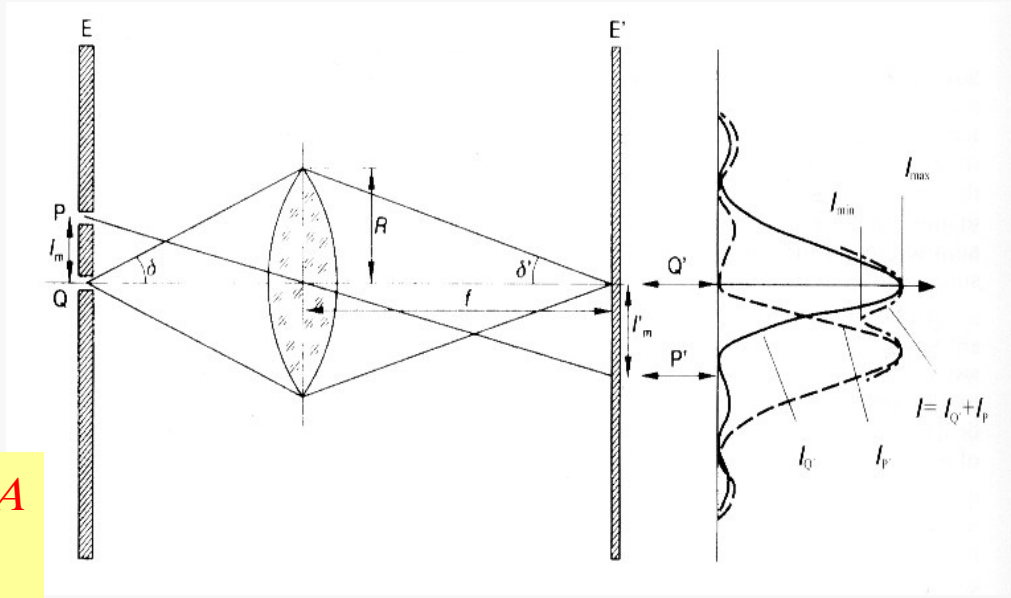
Figure 4 clarifies the connection between mask, diffraction and intensity distribution in the image plane. Due to diffraction two sharp features, P and Q, on the mask give rise to an overall intensity distribution on the sample. To resolve these two features the intensity distribution has to have a minimum between the two main maximums. It is useful to define the so called modulation transfer function (*MTF*) as:

$$MTF = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \qquad (6)$$

The higher the value – the higher the difference between the maximum and minimum intensity – the better the contrast between exposed and unexposed areas, the better is the resolution of the equipment. It should be noted that the *MTF* is only derived by properties of the optical system. It is a measure of the capabilities of the lithographic tool in printing structures.

Pattern contrast affected by optical diffraction

Contrast relevant in ruling photoresist impression



"Space resolution" ~ $k\,\lambda/NA$ with k ~ 0.5-0.9

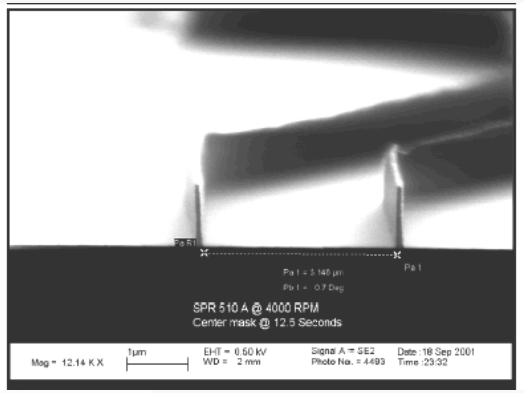
How can it be improved?

Da R. Waser Ed., Nanoelectronics and information technology (Wiley-VCH, 2003)

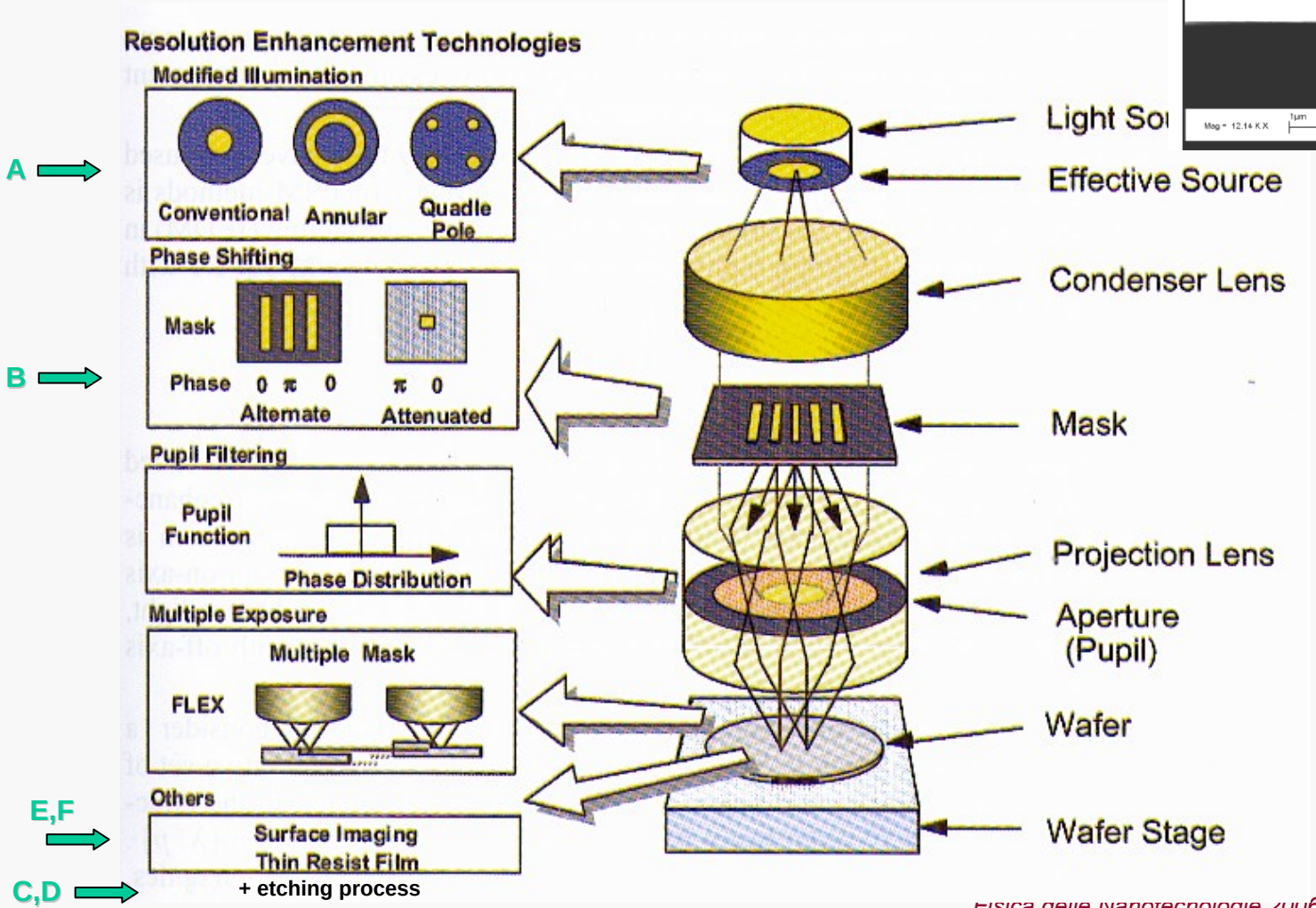
4. Strategies to improve resolution

Great efforts devoted in the past to push resolution down to the diffraction limit (or even slightly below)

Mostly, **optics** and **"chemistry"** are involved into



Example: features ~ 100 nm made with 350 nm radiation



A. Off-axis illumination

Off-Axis Illumination

To improve resolution without decreasing the wavelength or increasing *NA*, so-called off-axis illumination was applied. The method was already known as a contrast-enhancing technique for optical microscopes. With off-axis illumination, the light beam is directed from the mask towards the edge of the projection lens, and not, as in on-axis illumination, towards the center. In normal illumination with partially coherent light, there always is part of the light which is off-axis, but in the context here with off-axis illumination there is no on-axis component.

To understand the mode of operation of off-axis illumination, consider a line-and-spaces structure with pitch p. The incident light will be diffracted into a set of beams, of which only the undiffracted beam, the zero-order beam, travels in the direction of the incident light. The 1st order beam travels under the angle $|\theta_1| = \arcsin(\lambda / p)$. If p is too small, then $|\theta_{\pm 1}|$ is bigger than the acceptance angle α of the projection optics, then only the zero-order beam is projected to the sample (Figure 7a). But this does not carry any information of the pattern, and hence the pattern cannot be transferred onto the sample. At least the zero- and the 1st order beam have to be in the range of the aperture angle. If the incident light hits the mask under an angle $\Theta_0 < \alpha$ the undiffracted beam enters the projection lens at the edge, and the 1st order beam is still collected by the lens, and therefore a pattern transfer is still possible. The angle of incidence Θ_0 can be realized by inserting an aperture in the optical path between condenser and mask (Figure 7b).

Although the higher resolution is an advantage of off-axis illumination, the impact on the depth of focus (DOF) is of even greater value. In on-axis illumination, the beams of different deflection orders have to travel in different ways so they are phase-shifted to each other, which results in a lack of focus. In off-axis illumination, the zero order and 1st order beam reaches the projection lens at the same distance from the center, which means that their optical path length is the same. So the relative phase difference between these beams is zero, which increases the DOF dramatically.

Off-axis illumination is facilitated by an aperture (Figure 8) which is located in front of the condenser lens. It depends on the apertures shape which structures are improved. If there is an aperture as in Figure 8a, only the structures perpendicular to the arrangement of the apertures will be improved. The aperture shown in Figure 8b yields an improvement of structures which are adjusted to *good* angles – up/down or left/right direction. This is sufficient because in normal cases, the features are in a *good* arrangement. The aperture in Figure 8c even decreases this problem, but here the improvement in DOF is less.

When the resolution in principle has to be improved, then according to the Rayleigh criterion either the wavelength λ or the technology parameter k_1 have to be decreased, or the numerical aperture *NA* has to be increased.

Increasing *NA* means physically bigger lenses. Here the problem arises that it is difficult to produce huge lenses with the required quality; on the other hand the available materials also limit the physical size of the lenses. So there are still two possibilities of increasing the resolution: smaller λ and smaller k_1.

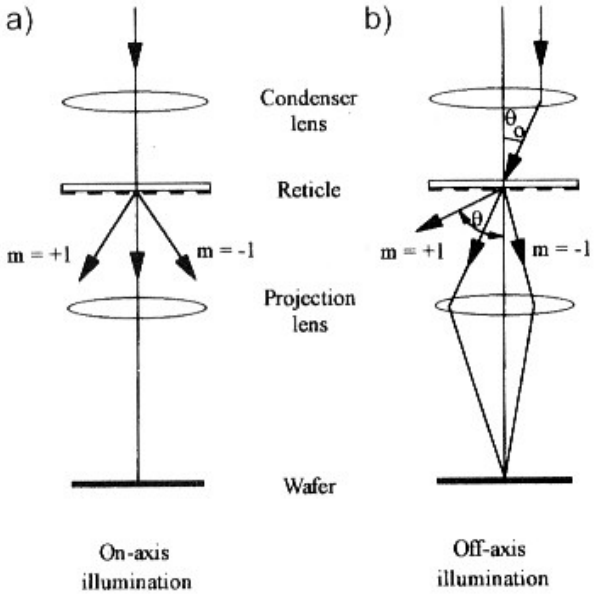


Figure 7:
(a) Optical path and deflection orders of on-axis and
(b) off-axis illumination. Note that with the same wavelength and structure size, the off-axis illumination allows the 1st order beam to pass the optical system [3]. A good description of off-axis illumination is also found in [6].

(Reminescent of confocal microscopy issues)

Da R. Waser Ed., Nanoelectronics and information technology (Wiley-VCH, 2003)

B. Phase-shifting masks I

so-called *Levenson* or *alternating phase shift masks* (PSM) can improve the resolution by 40 %. Unfortunately, this improvement is pattern-dependent; for a single structure there is no neighboring structure, so there is no light to interfere with. Even if there are structures which are not in a regular arrangement, there is no defined phase shift between these structures which could yield an improvement in the resolution of all structures.

The phase shift can be obtained by an additional transparent layer on the mask. If it has the refractive index n and thickness d, the phase shift is $\Phi = (n-1)2\pi d/\lambda$. So a shift of π is obtained, when the condition $d = \lambda/[2(n-1)]$ holds. On the other hand, it is also possible to recess the mask material so that the right optical path difference is obtained. But the etch depth can be controlled by the time only, and not, as in etching away an additional layer, by the thickness of the layer itself.

To deal with the drawbacks of alternating PSM, several other methods have been developed, which are described next. In rim-PSM, the whole mask is covered by a phase-shifter material and then with the resist. After development, the phase shifter is etched anisotropically and the masking layer is etched isotropically. By this a undercut under the phase shifter occurs at the rim of every structure. This also yields a resolution improvement, but not as much as with alternating PSM, although it is therefore not lim-ited to certain structures.

Retardation of light engineered to improve interference contrast (suitable for periodic or semiperiodic patterns)

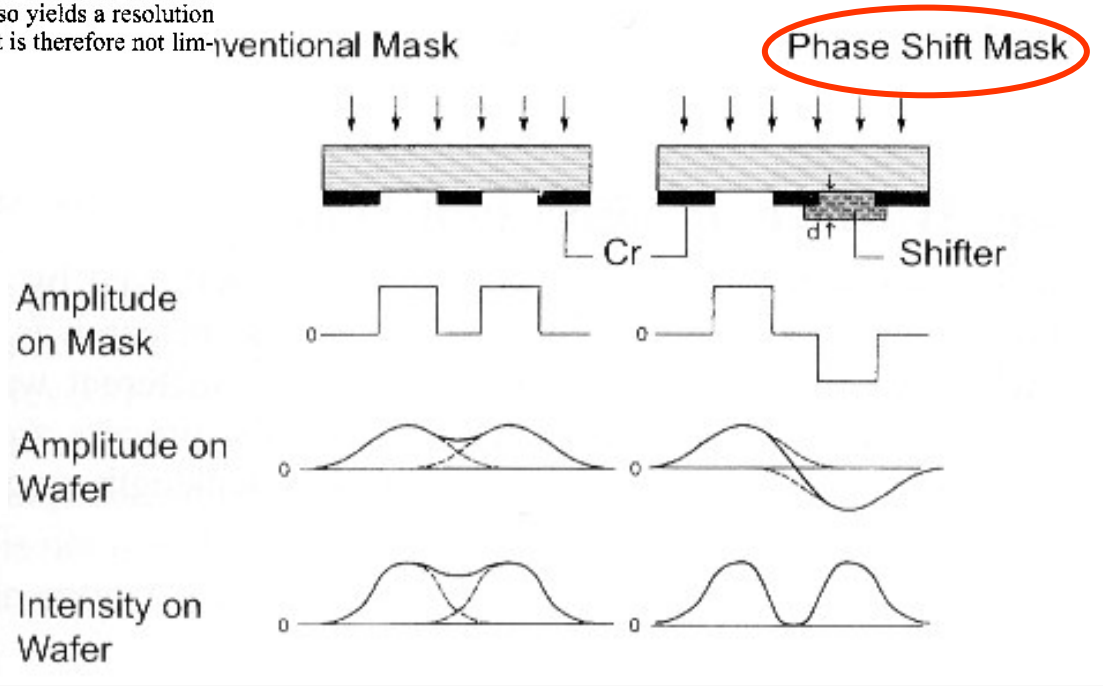


Figure 6: Comparison of the light amplitudes and intensities at the mask and on the wafer for a conventional and a phase shift mask. Note that the intensity on the wafer between the two features is zero for the phase shift mask [26].

Material and complexity issues

Reminders of light polarization

Phase-shifting masks act like half-waveplates to invert the (linear) polarization of the illumination beam

Waveplates operate by imparting unequal phase shifts to orthogonally polarized field components of an incident wave. This causes the conversion of one polarization state into another.

There are two cases. With linear birefringence, the index of refraction and hence phase shift differs for two orthogonally polarized linear polarization states. This is the operation mode of standard waveplates.

With circular birefringence, the index of refraction and hence phase shift differs for left and right circularly polarized components. This is the operation mode of polarization rotators.

d Waveplates: Linear Birefringence

Suppose a waveplate made from a uniaxial material has light propagating perpendicular to the optic axis. This makes the field component parallel to the optic axis an extraordinary wave and the component perpendicular to the optic axis an ordinary wave. If the crystal is positive uniaxial, $n_e > n_o$, then the optic axis is called the slow axis, which is the case for crystal quartz. For negative uniaxial crystals, $n_e < n_o$, the optic axis is called the fast axis.

The equation for the transmitted field E_2, in terms of the incident field E_1 is:

$$E_2 = s(s \bullet E_1)e^{i\phi_s} + f(f \bullet E_1)e^{i\phi_f}$$

where s and f are unit vectors along the slow and fast axes. This equation shows explicity how the waveplate acts on the field. Reading from left to right, the waveplate takes the component of the input field along its slow axis and appends the slow axis phase shift to it. It does a similar operation to the fast component.

The slow and fast axis phase shifts are given by:

$$\phi_s = n_s(\omega)\omega t / c = 2\pi n_s(\lambda)t / \lambda$$
$$\phi_f = n_f(\omega)\omega t / c = 2\pi n_f(\lambda)t / \lambda$$

where n_s and n_f are, respectively, the indices of refraction along the slow and fast axes, and t is the thickness of the waveplate.

To further analyze the effect of a waveplate, we throw away a phase factor lost in measuring intensity, and assign the entire phase delay to the slow axis:

$$E_2 = s(s \bullet E_1)e^{i\phi} + f(f \bullet E_1)$$
$$\phi = \phi_s - \phi_f = 2\pi(n_s(\lambda) - n_f(\lambda))t / \lambda$$
$$= 2\pi\Delta n(\lambda)t / \lambda$$

In the above, $\Delta n(\lambda)$ is the birefringence $n_s(\lambda) - n_f(\lambda)$. The dispersion of the birefringence is very important in waveplate design; a quarter waveplate at a given wavelength is never exactly a half waveplate at half that wavelength.

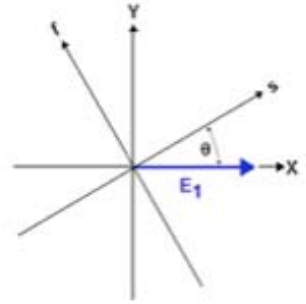
Let E_1 be initially polarized along X, and let the waveplate slow axis make an angle ζ with the X-axis. This orientation is shown in the figure below.

For a full waveplate:
$\phi = 2m\pi$, $T_{||} = 1$, and $T\perp = 0$, regardless of waveplate orientation.

For a half waveplate:
$\phi = (2m + 1)\pi$, $T_{||} = \cos^2 2\theta$, and $T\perp = \sin^2 2\theta$.

This transmission result is the same as if an initial linearly polarized wave were rotated through an angle 2θ. Thus, a half waveplate finds use as a polarization rotator.

For a quarter waveplate:
$\phi = (2m + 1)\pi/2$; i.e. an odd multiple of $\pi/2$. To analyze this, we have to go back to the field equation. Assume that the slow and fast axis unit vectors s and f form a right-handed coordinate system such that s x f = +z, the direction of propagation.



When the waveplate is placed between parallel and perpendicular polarizers the transmissions are given by:

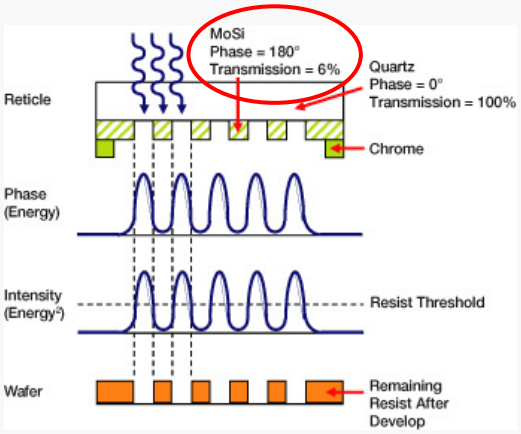
$$T_{||} \propto |E_{2x}|^2 = 1 - \sin^2 2\theta \sin^2 \phi/2$$
$$T\perp \propto |E_{2y}|^2 = \sin^2 2\theta \sin^2 \phi/2$$

Note that θ is only a function of the waveplate orientation, and f is only a function of the wavelength, the birefringence is a function of wavelength and the plate thickness.

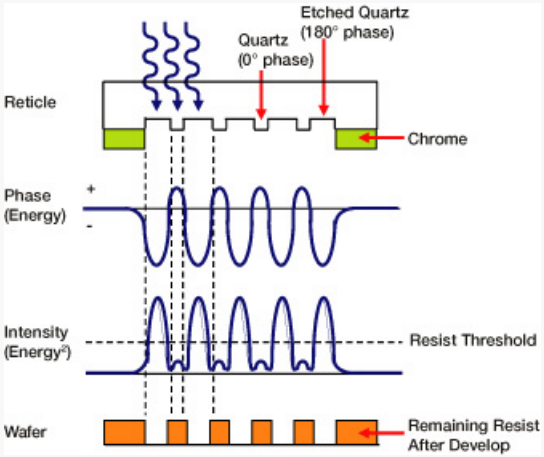
Phase Shift	Input Field Along (s+f)/√2	Input Field Along (s-f)/√2
$\phi = \pi/2 + 2m\pi$	RCP	LCP
$\phi = 3\pi/2 + 2m\pi$	LCP	RCP

Phase-shifting masks II

Attenuated Phase Shift Masks (AttPSM) form their patterns through adjacent areas of quartz and, for example, molybdenum silicide (MoSi). Unlike chrome, MoSi allows a small percentage of the light to pass through (typically 6% or 18%). However, the thickness of the MoSi is chosen so that the light that does pass through is 180° out of phase with the light that passes through the neighboring clear quartz areas. The light that passes through the MoSi areas is too weak to expose the resist, however the phase delta serves to "push" the intensity down to be "darker" than similar features in chrome. The result is a sharper intensity profile which allows smaller features to be printed on the wafer.



AltPSMs employ alternating areas of chrome and 180 degree-shifted quartz to form features on the wafer. AltPSM is a powerful but complex technology. The process of manufacturing the mask is considerably more demanding and expensive than that for Binary masks. Furthermore, the AltPSM must be accompanied by a second "Trim" mask, resulting in extra cost and decreased stepper throughput.

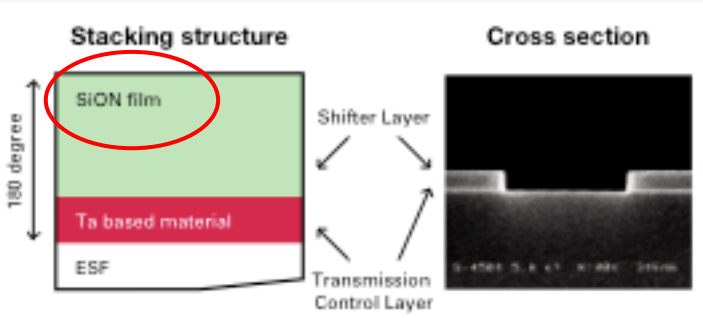


Attenuated PSM (att. PSM)

Silicon-oxinitride on metal (bilayer structure).

SiON shifter film was firstly developed by HOYA, and is now going into production as a co-work between HOYA and DNP.

This shifter will be used as our standard material for 157nm att. PSM as well as high transmission 193nm.



http://www.asml.com
http://www.dnp.com

Diligenti et al.
APL 75 489 (1999)

C. Anisotropic etching

An (110) silicon substrate (p doped, resistivity $=1-10\,\Omega\times$cm) was oxidized (1050 °C for 30 min) to obtain a 350 nm SiO_2 layer. Standard photolithography was employed to pattern the oxide; the mask used is shown in Fig. 1(a). It consisted of an array of lines (width 4 μm, length 3000 μm, spacing 4 μm) which must be aligned along the $[\bar{1}12]$ direction. After the oxide definition, an anisotropic etching was performed by means of an ethylenediamine-pyrocatechol (EDP) solution type F (fast) at 115 °C.[16] After 1 h of etching walls with an high aspect ratio [Fig. 1(b)] (height=35 μm, width in the range 1–4 μm, depending on the alignment accuracy along the $[\bar{1}12]$ direction) were obtained. For samples with a misalignment greater than 0.07 degrees there was a complete underetching of the planes.[17] Each array contained 1000 planes; Fig. 2 (top panel) shows a scanning electron microscope (SEM) micrograph of a cross section of an array of planes.

The samples were then etched in buffered HF (BHF) to remove the oxide mask layer indicated as SiO_2 in Fig. 1(b), and underwent a oxidation. The sequence of oxidation/etching steps allowed to reduce the wall thickness in a controlled way, and PL measurements were carried out after each step to investigate the dependence of the emission features on the wall width. Figure 2 (bottom panel) is a closer view of a cross section of the planes, which shows that the Si core thickness is not uniform, an effect probably due to a minor oxygen diffusion at the bottom of the walls.

Etching can lead to reduce the feature size (not the pitch, though)

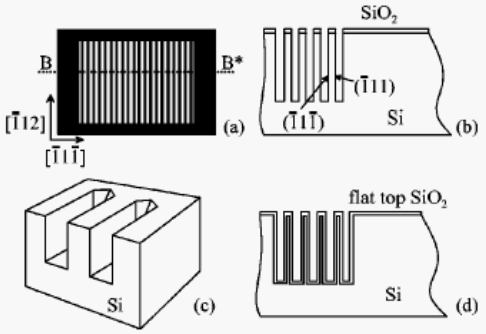
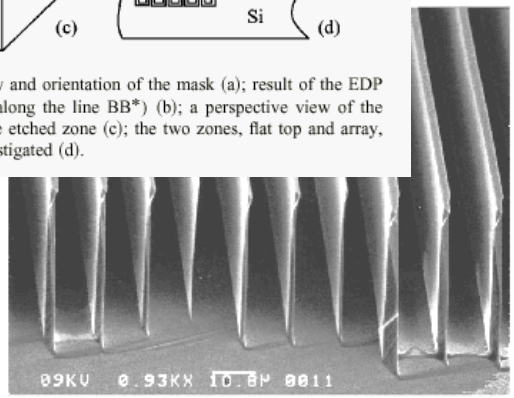
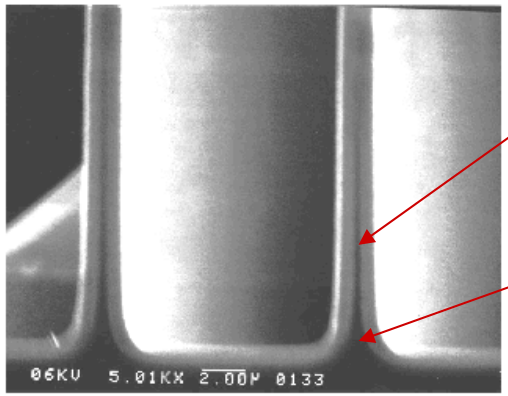


FIG. 1. Schematic view and orientation of the mask (a); result of the EDP etching (cross section along the line BB*) (b); a perspective view of the planes at the end of the etched zone (c); the two zones, flat top and array, where the PL was investigated (d).

Mask alignment





Silicon oxide

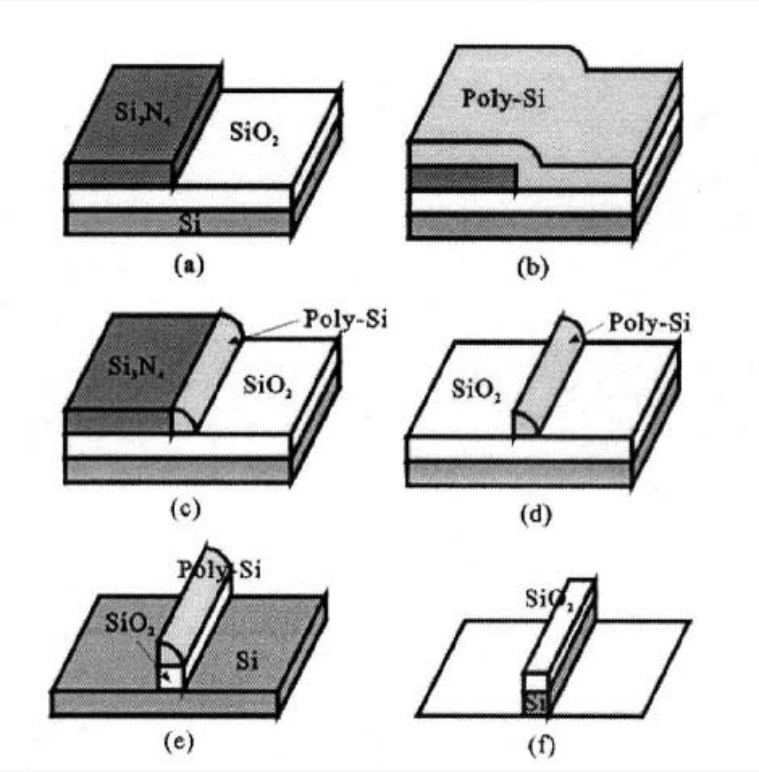
Crystalline Silicon core

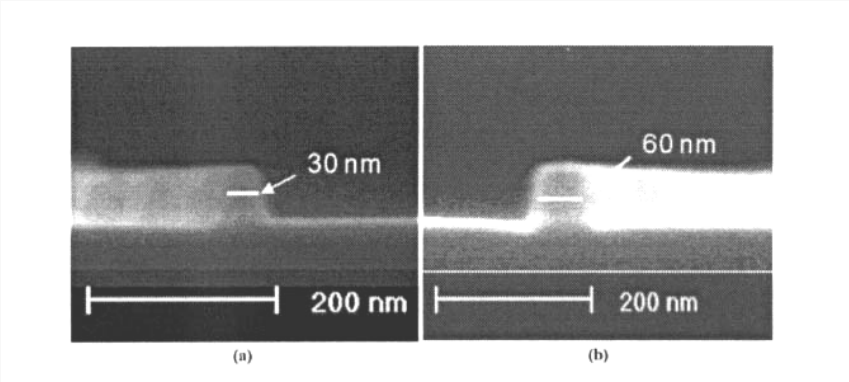
FIG. 2. SEM view of the array cross section (top panel); a closer view showing the oxidation effect (bottom panel).

D. Side wall patterning

Processo di fabbricazione “complesso”
(ma economico!!) per creare **Si-nanowires**:

a) ossidazione dry (spess. ≤ 0.5 μm):
$Si\ (s) + O_2 \rightarrow SiO_2$
seguita da deposizione Si_3N_4 e patterning via lito. ottica convenzionale
b) CVD poly-Si (pirolisi SiH_4 a bassa p):
SiH_4 + calore(T~1100 °C) $\rightarrow$ Si (s) +2 H_2 (g)
c) Reactive Ion Etching del poly-Si con fascio ionico “inclinato”
--> rimane poly-Si solo sui bordi
d) wet chemical etching selettivo (soprattutto nitruro) con H_3PO_4
e) rimozione ossido (poly-Si funge da maschera) con etching selettivo
f) rimozione nitruro con RIE non inclinata





SEM cross sections (in “prospettiva”)

(Sometimes) by playing with etching process, results can be achieved comparable to electron beam lithography (once more, feature size, not pitch)

E. Interference (holographic) lithography

The mask is replaced by an interference pattern

Feature size can be pushed well below diffraction limit

JM Carter et al., MIT Annual Report 2003

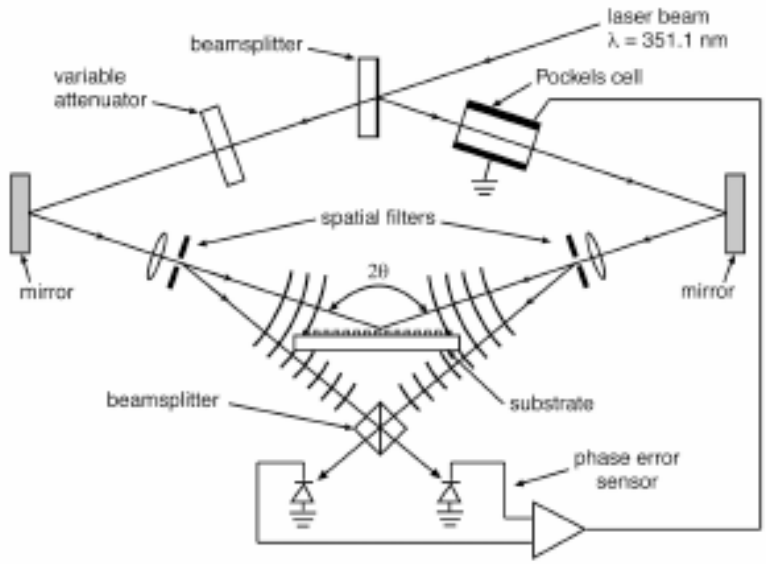


Fig. 28: Schematic of one of the MIT interferometric lithography systems. This system occupies a 2x3m optical bench in a class 100 clean environment. The beamsplitter directs portions of the two interfering spherical beams to photodiodes. A feedback locking is achieved by differentially amplifying the photodiode signals and applying a correction to the Pockels cell which phase shifts one of the beams in order to stabilize the standing wave pattern at the substrate.

Pitch improved but
only space-regular patterns can be achieved
and covered area is typically small

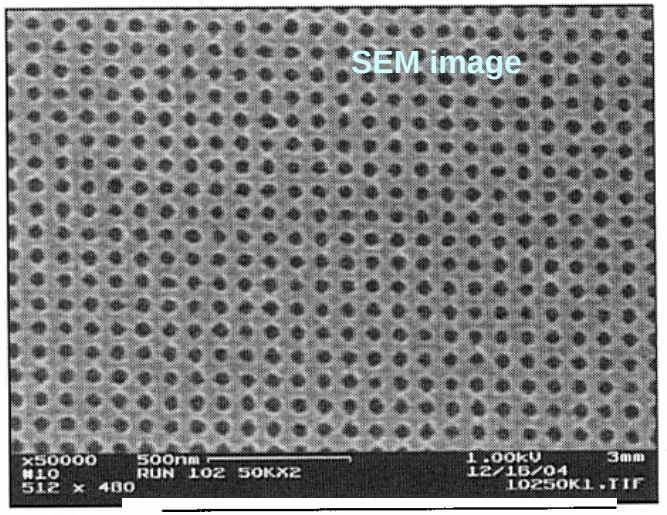


Figure 2. Arrays of 45-nm-diameter holes patterned on a 90-nm pitch with "dry" 157-nm interference lithography. The photoresist is negative-tone-acting hydrogen silsesquioxane (HSQ). The pattern was obtained with two crossed exposures. HSQ is a SiO_2-like material that acts as an effective growth mask in the epitaxy of III–V crystals. Therefore, the pattern in this figure can be employed as a template for homo- and heteroepitaxy of nanoscale structures, without the need for pattern transfer. Size marker is 500 nm.

MRS Bull., vol. 30 (dec. 2005)

F. Immersion lithography

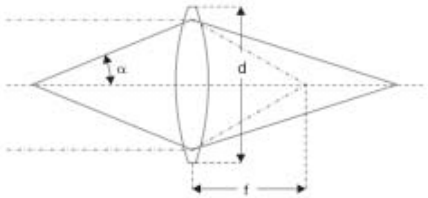


Figure 3. Numerical aperture.

The medium between the lens and the wafer being exposed needs to have an index of refraction >1, have low optical absorption at 193nm, be compatible with photoresist and the lens material, be uniform and non-contaminating. Surprisingly, ultrapure water may meet all of these requirements. Water has an index of refraction $n \cong 1.47$ [3], absorption of <5% at working distances of up to 6mm [4], is compatible with photoresist and lens and in it's ultrapure form is non-contaminating. There have been observations of significant variations in absorption between ultrapure water samples [4], but this is probably due to contaminants or dissolved gases. Plugging in $n = 1.47$ into equation 3 and assuming $\sin\alpha$ can reach 0.93, then the resolution limits for 193nm immersion lithography are

$$W = \frac{k_1\lambda}{n\sin\alpha} = \frac{0.25 \times 193}{1.47 \times 0.93} = 35nm \quad (4)$$

Resolution is inversely proportional to *n*

35nm theoretical resolution carries 193nm exposure beyond 2007. Similar techniques applied to 157nm exposure could carry optical lithography even further, although it should be noted that water is not a usable medium at 157nm and suitable mediums are still being researched.

4.0. Immersion lithography systems issues

There are a number of practical issues to implementing immersion lithography that still need to be resolved. The stage on a 193nm exposure tool steps from location to location across the wafer scanning the reticle image for each field. In order to achieve high throughput the stage must accelerate rapidly, move accurately to the next field location, settle, scan the image and then step to the next location all in a short period of time. Maintaining a consistent bubble free liquid between the lens and the wafer is very difficult. There are basically three approaches to the problem.

The first approach is to submerge the whole chuck, wafer and lens in a pool of water. The issue with the pool approach is that a complex system of servo motors and laser interferometers are required to accurately move the chuck, see figure 4, and submerging this whole system would require significant engineering.

The second approach is to limit the pool size to the top of the chuck. This technique would keep all of the chuck control mechanisms out of the water but would add considerable mass to the chuck that must rapidly accelerate.

The third technique and most likely to be used [4], is too dispense the water between the lens and the wafer with a nozzle and rely on surface tension to maintain a "puddle". Figure 5 illustrates the puddle approach. Note how the wafer is recessed into the chuck in figure 5. The lip around the chuck is to allow edge the puddle to extend off the edge of the wafer during edge die exposure.

One issue that is likely to be significant for immersion lithography is temperature control. Variations in temperature cause variations in n and therefore image distortion. Maintaining temperature uniformity with a rapidly moving stage and a pulsed laser passing through the fluid will likely be a significant challenge.

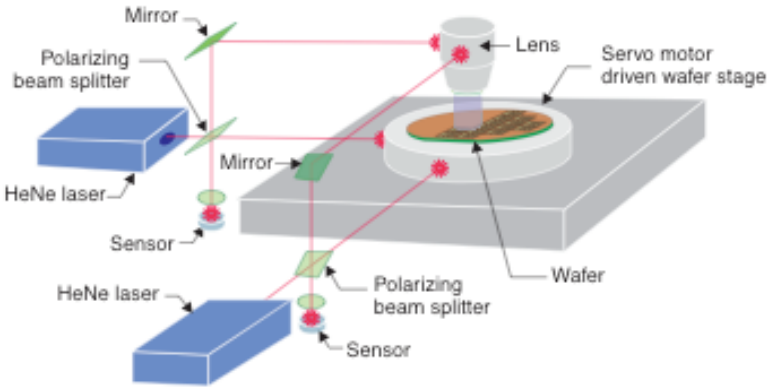


Figure 4. Stepping exposure system stage control

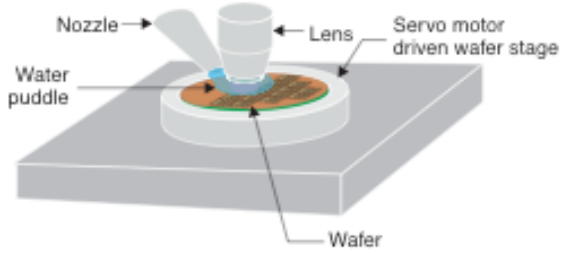


Figure 5. Immersion lithography. Stage control omitted for clarity.

***NA* can be improved *but* technical and material issues exist (thin liquid layer, absorption of light, decomposition, …)**

Light sources for optical lithography

Rather obvious recipe to further improve space resolution: decrease the wavelength

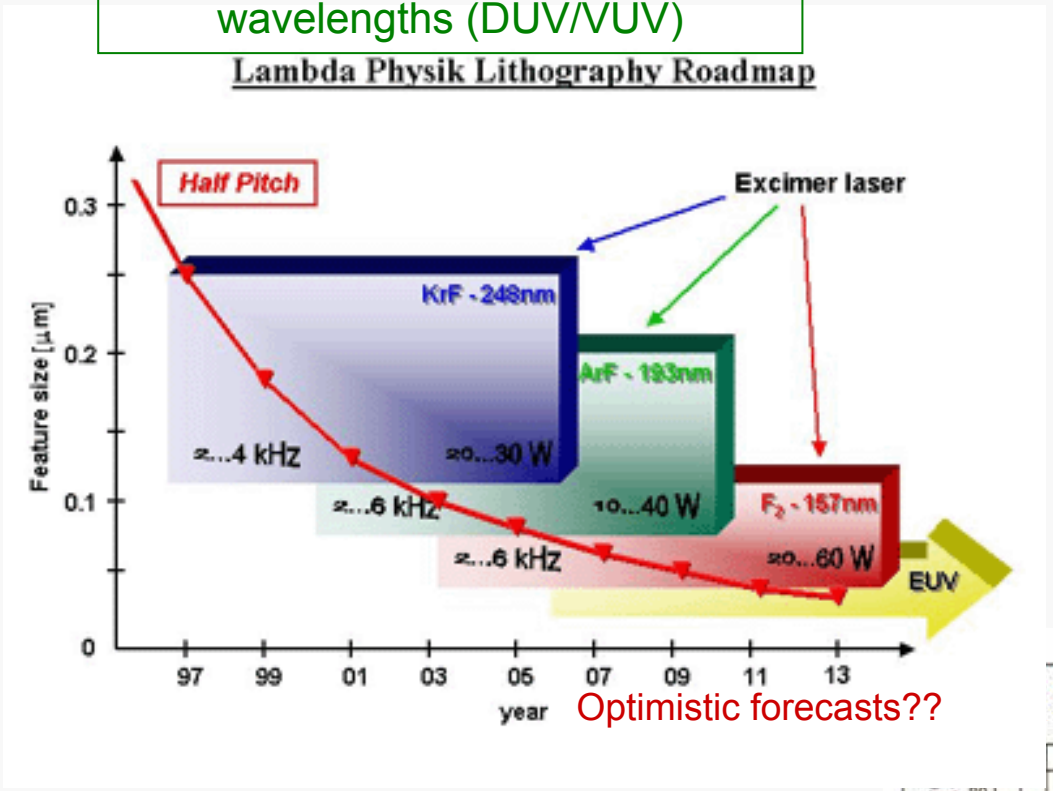
Wavelength [nm]	Source	Range
436	Hg arc lamp	G-line
405	Hg arc lamp	H-line
365	Hg arc lamp	I-line
248	Hg/Xe arc lamp; KrF excimer laser	Deep UV (DUV)
193	ArF excimer laser	DUV
157	F_2 laser	Vacuum UV (VUV)
~10	Laser-produced plasma sources	Extreme UV (EUV)
~1	X-ray tube; synchrotron	X-ray

EXCIMER LASERS: **lasers with an electric discharge plasma as gain medium, in which optical gain is generated by excited dimers (or other molecules) with an anti-binding electronic ground state**

An excimer laser is a powerful kind of laser which is basically always operated in the ultraviolet (UV) spectral region and generates nanosecond pulses. The excimer gain medium is a gas mixture, typically containing a noble gas (rare gas) (e.g. argon, krypton, or xenon) and a halogen (e.g. fluorine or chlorine, e.g. as HCl), apart from helium and/or neon as buffer gas. An excimer gain medium is pumped with short (nanosecond) current pulses in an electric discharge (or sometimes with an electron beam), which create so-called excimers (excited dimers) – molecules which represent a bound state of their constituents only in the excited electronic state, but not in the electronic ground state. (Precisely speaking, a dimer is a molecule consisting of two equal molecules, but the term excimer is nowadays understood to include asymmetric molecules such as XeCl as well. The term rare gas halide lasers would actually be more appropriate, and the term exciplex laser is sometimes used.) After stimulated or spontaneous emission, the excimer rapidly dissociates, so that reabsorption of the generated radiation is avoided. This makes it possible to achieve a rather high gain.
Typical excimer lasers emit pulses with a repetition rate up to a few kilohertz and average output powers between a few watts and hundreds of watts, which makes them the most powerful laser sources in the ultraviolet region, particularly for wavelengths below 300 nm. The power efficiency varies between 0.2% and 2%.

Search for shorter (radiation) wavelengths

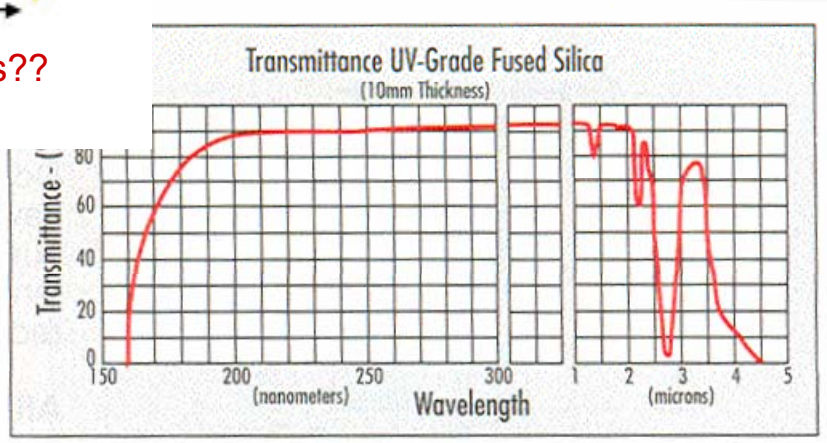
Continuous development of laser sources with smaller and smaller wavelengths (DUV/VUV)



Optimistic forecasts??

Not all wavelengths are accessible!!

Excimer	Wavelength
F2	157 nm
ArF	193 nm
KrF	248 nm
XeBr	282 nm
XeCl	308 nm
XeF	351 nm



VUV implies:

Material absorption issues

Radiation handling issues

5. X-ray lithography (XRL)

XR source of choice: synchrotron
Brilliant XR beam
--> proximity mode masks

Resist: typ. PMMA
(critical sensitivity
--> large dose, ~ 2 J/cm^2)

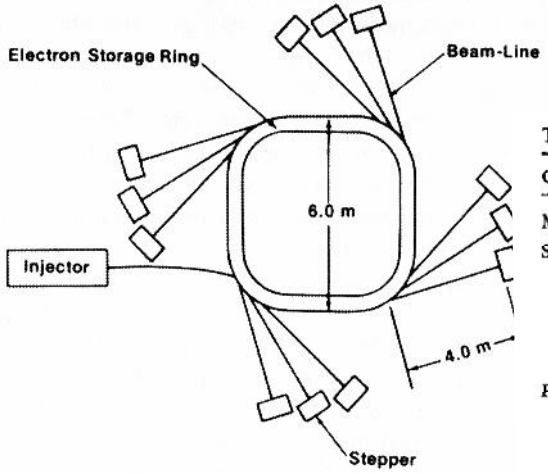
Masks:
Typ. Si membranes

Diffraction problems *virtually* removed thanks to the very short wavelength

Synchrotron

Synchrotron Radiation Parameters

Beam current	200–300 mA
Energy	0.6–1.4 GeV
Critical wavelength (l_c)	1–20 Å
Beam lifetime	5–12 hr
Injection energy	50 MeV full energy



Da M. Madou,
Fundamentals of microfab.,
CRC (1997)

TABLE 1.5 Optical vs. X-Ray Mask

Optical Mask	X-Ray Mask
Mask design: CAD	Mask design: CAD
Substrate preparation	Substrate preparation
Quartz	Thin membrane substrate (Si, Be, Ti, ...)
Thin metal film deposition	Deposit plating base (50 Å Cr, then 300 Å Au)
Pattern delineation	Pattern delineation
Coat substrate with resist	Coat with resist
Expose pattern (optical, e-beam)	Expose pattern (optical, e-beam)
Develop pattern etch Cr layer	Develop pattern
Strip resist	Absorber definition:
	Electroplate Au (~15 μm for hard X-rays)
	Strip resist
Cost: $1K–$3K	Cost: $4K–$12K
Duration: 3 days	Duration: 10 days

Effective resolution ~ tens of nm
Large depth of field -> suitable for high aspect-ratio features, micromachining , …

Technological limits of XRL

4 X-Ray Lithography

Decreasing the wavelength even further into the x-ray range yields so-called x-ra lithography. For these short wavelengths it is not possible to set up an optical path nei ther in reflection optics nor in refraction optics. On one hand, there is no material whicl is transparent enough to make lenses or masks from, and, on the other hand, it is not pos sible to make Bragg-reflectors. The individual layers in the layer stack have to have thickness of $\lambda/4$, which corresponds to a layer thickness of ~0.3 nm. This is in the rang of the thickness of one monolayer and is not achievable..

Projection x-ray-lithography is therefore not possible, but proximity x-ray lithogra phy (PXL) is possible. The advantages are the high resolution limit ($\sim\sqrt{\lambda\cdot(g+d)}$ which is about 30 nm for 1 nm exposure wavelength) and the insensitivity to organi contamination. These contaminations (as all low atomic number materials) do no absorb the x-rays, and hence are not printed onto the sample.

But there are some limitations. Consider a source with diameter a of 1 mm at dis tance L of 1 m towards the mask and a proximity gap g of 10 µm. Then there is th so-called *penumbral blur* $\xi = a \cdot g / L \sim 10$ nm, which limits the resolution (Figure 16) Furthermore, the pattern is not transferred correctly to the sample. Even if a point sourc is used, there is a displacement Δ of $\Delta = r \cdot g / L$, where r is the radial position on th sample (Figure 16). This error can be eliminated if it is taken into account when th mask pattern is generated.

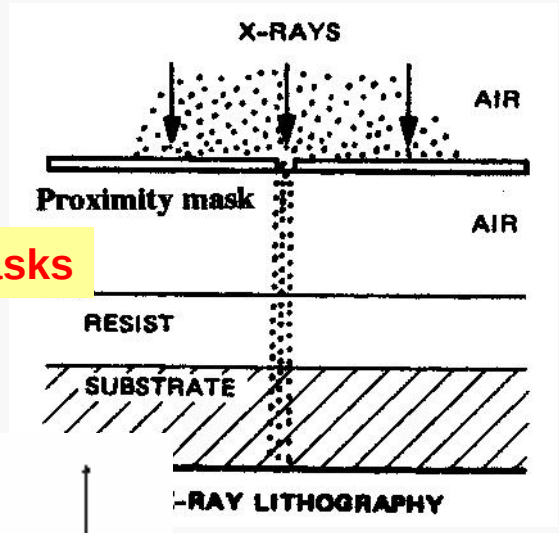
Nevertheless, if synchrotron radiation is used, a high intense beam of parallel ligh is available so these errors do not occur. This parallel beam has another advantage: Du to the small deflection the exposure shows a high depth of focus of several µm, facilitat ing exposures of textured substrates or of thick resists (Figure 17).

The problem for PXL is the masks. Since there is no material which is as transpar ent to x-ray as quartz to DUV, the carrier layer has to be thin (1 – 2 µm). On the othe hand, there is also no material which is as opaque to x-ray as chromium to DUV, so th masking layer has to be thick enough (300 – 500 nm). A carrier layer of 1 µm SiC onl has a transparency of 57 %, while a masking layer of Au still lets 14 % of the light pass The absorbed light will heat the mask so that it expands, which leads to another uncer tainty in the pattern transfer. Furthermore, PXL is a non-reduction printing method, s the features on the mask are of the same size as on the sample. This makes the produc tion of the masks very complicated when the target ist the sub-100 nm range.

The mask production sequence is as follows: On a silicon wafer, a thin membran layer is deposited (e.g. SiC, Si_3N_4). Onto this layer, a chromium etch stop layer and th masking layer of 300 – 500 nm of a high-atomic number material is evaporated (e.g. Au Ta). Then the mask is coated with an e-beam resist and exposed in an e-beam direct-write system. The resist is used to etch the masking layer with an etch stop on th chromium so the membrane is not hurt.

The commonly used DUV resists show good process aptitude.

Proximity masks



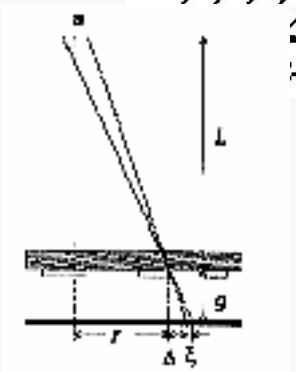


Figure 16: Penumbral blur ξ and displacement error Δ for proximity x-ray lithography. L is the distance from source to mask, g is the proximity gap and a is the lateral diameter of the source [11].

Penumbral blur

Mask materials

Large scattering of secondary electrons limits the resolution

Conclusions

✓ Optical lithography has been the dominant technique in microelectronics and still plays a relevant role (*this decade, at least)*

✓ Lithography and microscopy are “complementary” techniques sharing common problems

✓ Fundamental limitations exist preventing its exploitation in the nanotechnology realm *(problems for the next decades)*

✓ Efforts are being devoted to keep the pace of miniaturization *(as far as possible)*

✓ Technical complexity is growing up, technological limitations (e.g., materials) appear

✓ *Alternative nanofabrication approaches urgently needed!!*