METHOD AND APPARATUS FOR ANGULAR-RESOLVED SPECTROSCOPIC LITHOGRAPHY CHARACTERIZATION

FIELD

[0001] The present invention relates to methods of inspection usable, for example, in the manufacture of devices by lithographic techniques and to methods of manufacturing devices using lithographic techniques.

BACKGROUND

[0002] In a manufacturing process using a lithographic projection apparatus, a pattern (e.g. in a mask) is imaged onto a substrate that is at least partially covered by a layer of radiation-sensitive material (resist) by the changes of either optical properties or surface physical properties of the resist. Alternatively, the imaging step may use a resistless process such as etched grating or nano-imprint technology. Prior to this imaging step, the substrate may undergo various procedures, such as priming, resist coating and a soft bake. After exposure, the substrate may be subjected to other procedures, such as a post-exposure bake (PEB), development, a hard bake and measurement/inspection of the imaged features. This array of procedures is used as a basis to pattern an individual layer of a device, e.g. an IC. Such a patterned layer may then undergo various processes such as etching, ion-implantation (doping), metallization, oxidation, chemical-mechanical polishing, etc., all intended to finish off an individual layer. If several layers are required, then the whole procedure, or a variant thereof, will have to be repeated for each new layer. Eventually, an array of devices will be present on the substrate (wafer). These devices are then separated from one another by a technique such as dicing or sawing, whence the individual devices can be mounted on a carrier, connected to pins, etc.

[0003] The measurement and inspection step after development of the resist (or substrate surface in the case of etching), referred to as in-line because it is carried out in the normal course of processing production substrates, typically serves two purposes. Firstly, it is desirable to detect any target areas where the pattern in the developed resist is faulty. If a sufficient number of target areas are faulty, the substrate can be stripped of the patterned resist and re-exposed, hopefully correctly, rather than making the fault permanent by carrying out a process step, *e.g.*, an etch, with a faulty pattern. Secondly, the measurements

may allow errors in the lithographic apparatus, *e.g.* illumination settings or exposure dose, to be detected and corrected for in subsequent exposures. However, many errors in the lithographic apparatus cannot easily be detected or quantified from the patterns printed in resist. Detection of a fault does not always lead directly to its cause. Thus, a variety of off-line procedures for detecting and measuring errors in the lithographic apparatus are known. These may involve replacing the substrate with a measuring device or carrying out exposures of special test patterns, *e.g.*, at a variety of different machine settings. Such off-line techniques take time, often a considerable amount, during which the end products of the apparatus will be of an unknown quality until the measurement results are made available. Therefore, in-line techniques, ones which can be carried out at the same time as production exposures, for detecting and measuring errors in the lithographic apparatus, are usually preferred.

[0004] Scatterometry is one example of an optical metrology technique that can be used for in-line measurements of CD and overlay. There are two main scatterometry techniques:

[0005] (1) Spectroscopic scatterometry measures the properties of scattered light at a fixed angle as a function of wavelength, usually using a broadband light source such as xenon, deuterium, or halogen based light source such as a xenon arc lamp. The fixed angle can be normally incident or obliquely incident.

[0006] (2) Angle-resolved scatterometry measures the properties of scattered light at a fixed wavelength as a function of angle of incidence, usually using a laser as a single wavelength light source.

[0007] The structure giving rise to a reflected spectrum is reconstructed, *e.g.*, using real-time regression or by comparison to a library of patterns derived by simulation. Reconstruction involves minimization of a cost function. Both approaches calculate the scattering of light by periodic structures. The most common technique is Rigorous Coupled-Wave Analysis (RCWA), though light scattering can also be calculated by other techniques such as Finite Difference Time Domain (FDTD) or Integral Equation techniques.

[0008] A problem with known angle resolved scatterometry techniques is that they only detect one wavelength at a time so spectra with more than one wavelength have to have those wavelengths time-multiplexed, which increases the total acquisition time taken

to detect and process the spectra. In spectroscopic scatterometry, an extended light source with a large etendue is used. Since a small grating must be illuminated with a small spread in angle of incidence, a lot of light from this extended source is wasted. This results in low light levels on the detector that lead to long acquisition times, which have a negative impact on throughput. If short acquisition times are chosen, the measurement results might not be stable.

SUMMARY

[0009] Accordingly, it would be advantageous, for example, to provide a method of measuring overlay and grating shape parameters (such as grating asymmetry and alignment) during manufacture of devices using lithographic techniques and measurement of an angle resolved spectrum in a pupil plane (or back focal plane) of a high NA (numerical aperture) lens. Projection system aberrations, etc. can also be measured in order to be corrected or compensated for.

[0010] Embodiments of the present invention may encompass hardware that is capable of measuring angle-resolved spectra at multiple wavelengths simultaneously, of carrying out immersion scatterometry and a focus measurement method for an angle resolved scatterometer, and of measuring intensity noise of a radiation source with a 2-D detector array. Furthermore, embodiments of the present invention may encompass applications of the hardware including measuring overlay through the measurement of asymmetry of scattered light and measuring small line shape variations via Rayleigh anomalies and high diffraction orders of scattered light.

[0011] Although specific reference may be made in this text to the use of the apparatus according to the invention in the manufacture of ICs, it should be explicitly understood that such an apparatus has many other possible applications. For example, it may be employed in the manufacture of integrated optical systems, guidance and detection patterns for magnetic domain memories, liquid-crystal display panels, thin-film magnetic heads, etc. The skilled artisan will appreciate that, in the context of such alternative applications, any use of the terms "reticle", "wafer" or "die" in this text should be considered as being replaced by the more general terms "mask", "substrate" and "target portion", respectively.

[0012] In the present document, the terms "radiation" and "beam" are used to encompass all types of electromagnetic radiation, including ultraviolet radiation (*e.g.* with a wavelength of 365, 248, 193, 157 or 126 nm) and EUV (extreme ultra-violet radiation, *e.g.*, having a wavelength in the range 5-20 nm), as well as particle beams, such as ion beams or electron beams.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] Embodiments of the invention will now be described, by way of example only, with reference to the accompanying schematic drawings in which corresponding reference symbols indicate corresponding parts and in which:

[0014] Figure 1 depicts a lithographic projection apparatus that may be used to perform a method according to an embodiment of the invention;

[0015] Figure 2 depicts a scatterometer;

[0016] Figure 3 depicts the general operating principle of measuring an angle resolved spectrum in the pupil plane of a high NA lens according to an embodiment of the invention;

[0017] Figures 4a and 4b depict the use of an embodiment of the present invention in determining overlay;

[0018] Figure 5 depicts the use of a non-polarizing beam splitter for coupling off a portion of a radiation beam according to an embodiment of the invention;

[0019] Figure 6 depicts a wavelength multiplexer according to an embodiment of the invention;

[0020] Figure 7 depicts a wavelength demultiplexer according to an embodiment of the invention;

[0021] Figure 8 depicts a knife edge at an intermediate object plane according to an embodiment of the invention;

[0022] Figures 9a and 9b depict a shaped obscuration in an inspection beam according to an embodiment of the invention;

[0023] Figure 10 depicts a detected image of different diffraction orders of scattered spectra according to an embodiment of the invention;

[0024] Figure 11 depicts a scatterometer with two illumination spots according to an embodiment of the invention;

[0025] Figure 12 depicts an ellipsometer according to an embodiment of the present invention;

[0026] Figure 13 depicts a scatterometer for detecting images in the pupil plane and the image plane according to an embodiment of the present invention; and

[0027] Figure 14 depicts an grating overlay of twice the pitch of a grating.

DETAILED DESCRIPTION

[0028] Figure 1 schematically depicts a lithographic projection apparatus useable in a method according to an embodiment of the invention. The apparatus comprises:

[0029] · a radiation system Ex, IL, for supplying a projection beam PB of radiation (*e.g.* DUV radiation), which in this particular case also comprises a radiation source LA;

[0030] · a first object table (mask table) MT provided with a mask holder for holding a mask MA (*e.g.* a reticle), and connected to a first positioning device for accurately positioning the mask with respect to item PL;

[0031] · a second object table (substrate table) WT provided with a substrate holder for holding a substrate W (*e.g.* a resist-coated silicon wafer), and connected to a second positioning device for accurately positioning the substrate with respect to item PL;

[0032] · a projection system ("projection lens") PL (*e.g.* a refractive lens system) for imaging an irradiated portion of the mask MA onto a target portion C (*e.g.* comprising one or more dies) of the substrate W.

[0033] As here depicted, the apparatus is of a transmissive type (*e.g.* has a transmissive mask). However, in general, it may also be of a reflective type, for example (*e.g.* with a reflective mask). Alternatively, the apparatus may employ another kind of patterning device, such as a programmable mirror array of a type as referred to above.

[0034] The source LA (*e.g.* an excimer laser) produces a beam of radiation. This beam is fed into an illumination system (illuminator) IL, either directly or after having traversed conditioning means, such as a beam expander Ex, for example. The illuminator IL may comprise adjusting means AM for setting the outer and/or inner radial extent (commonly referred to as σ-outer and σ-inner, respectively) of the intensity distribution in the beam. In addition, it will generally comprise various other components, such as an

integrator IN and a condenser CO. In this way, the beam PB impinging on the mask MA has a desired uniformity and intensity distribution in its cross-section.

[0035] It should be noted with regard to Figure 1 that the source LA may be within the housing of the lithographic projection apparatus (as is often the case when the source LA is a mercury lamp, for example), but that it may also be remote from the lithographic projection apparatus, the radiation beam which it produces being led into the apparatus (e.g. with the aid of suitable directing mirrors); this latter scenario is often the case when the source LA is an excimer laser. The current invention and claims encompass both of these scenarios.

[0036] The beam PB subsequently intercepts the mask MA, which is held on a mask table MT. Having traversed the mask MA, the beam PB passes through the projection lens PL, which focuses the beam PB onto a target portion C of the substrate W. With the aid of the second positioning device (and an interferometric measuring device IF), the substrate table WT can be moved accurately, *e.g.* so as to position different target portions C in the path of the beam PB. Similarly, the first positioning device can be used to position the mask MA accurately with respect to the path of the beam PB, *e.g.* after mechanical retrieval of the mask MA from a mask library, or during a scan. In general, movement of the object tables MT, WT will be realized with the aid of a long-stroke module (coarse positioning) and a short-stroke module (fine positioning), which are not explicitly depicted in Figure 1. However, in the case of a stepper (as opposed to a step-and-scan apparatus) the mask table MT may just be connected to a short stroke actuator, or may be fixed.

[0037] The depicted apparatus can be used in two different modes:

[0038] 1. In step mode, the mask table MT is kept essentially stationary, and an entire mask image is projected at one time (*i.e.* a single "flash") onto a target portion C. The substrate table WT is then shifted in the X and/or Y directions so that a different target portion C can be irradiated by the beam PB;

[0039] 2. In scan mode, essentially the same scenario applies, except that a given target portion C is not exposed in a single "flash". Instead, the mask table MT is movable in a given direction (the so-called "scan direction", *e.g.* the Y direction) with a speed v, so that the projection beam PB is caused to scan over a mask image; concurrently, the substrate table WT is simultaneously moved in the same or opposite direction at a speed $V = Mv$, in which M is the magnification of the projection system PL (typically, $M =$

1/4 or 1/5). In this manner, a relatively large target portion C can be exposed, without having to compromise on resolution.

One or more properties of the surface of a substrate 6 may be determined using a scatterometer such as that depicted in Figure 2. In an embodiment, the scatterometer comprises a broadband (white light) radiation source 2, which directs radiation onto a substrate 6. An extended broadband radiation source may be configured to provide the radiation beam with a wavelength of at least 50nm to the substrate surface. The reflected radiation is passed to a spectrometer detector 4, which measures a spectrum 10 (intensity as a function of wavelength) of the specular reflected radiation. From this data, the structure or profile giving rise to the detected spectrum may be reconstructed, *e.g.* by Rigorous Coupled Wave Analysis and non-linear regression or by comparison with a library of simulated spectra as shown at the bottom of Figure 2. In general, for the reconstruction, the general form of the structure is known and some parameters are assumed from knowledge of the process by which the structure was made, leaving only a few parameters of the structure to be determined from the scatterometry data.

[0040] The scatterometer may be a normal-incidence scatterometer or an oblique-incidence scatterometer. Variants of scatterometry may also be used in which the reflection is measured at a range of angles of a single wavelength, rather than the reflection at a single angle of a range of wavelengths.

[0041] In one or more embodiments described below, there is used a scatterometer configured to measuring a property of a substrate by measuring, in a pupil plane 40 of a high NA lens, a property of an angle-resolved spectrum reflected from the substrate surface 6 at a plurality of angles and wavelengths as shown in Figure 3. The scatterometer comprises a radiation source 2 configured to project radiation onto the substrate and a detector 32 configured to detect the reflected spectra. The pupil plane is the plane in which the radial position of radiation defines the angle of incidence and the angular position defines the azimuth angle of the radiation and any substantially conjugate plane. The detector 32 is placed in the pupil plane of the high NA lens. The NA is high and, in an embodiment, at least 0.9 or at least 0.95. Immersion scatterometers may even have lenses with an NA over 1.

[0042] Previous angle-resolved scatterometers have only measured the intensity of scattered light. An embodiment of the present invention allows several wavelengths to be

measured simultaneously at a range of angles. The properties measured by the scatterometer for different wavelengths and angles may include the intensity of transverse magnetic (TM) and transverse electric (TE) polarized light and the phase difference between the TM and TE polarized light.

[0043] Using a broadband light source (*i.e.* one with a wide range of light frequencies or wavelengths – and therefore of colors) is possible, which gives a large etendue, allowing the mixing of multiple wavelengths. The plurality of wavelengths in the broadband light, in an embodiment, each has a bandwidth of, say, $\delta\lambda$ and a spacing, therefore, of at least 2 $\delta\lambda$ (*i.e.* twice the wavelength). Several "sources" of radiation can be different portions of an extended radiation source that have been split using, say, fiber bundles. In this way, angle resolved scatter spectra can be measured at multiple wavelengths in parallel. A 3-D spectrum (wavelength and two different angles) may be measured, which contains more information than a 2-D spectrum. This allows more information to be measured which increases metrology process robustness.

[0044] A scatterometer of an embodiment of the present invention is shown in Figure 3. The light source 2 is focused using lens system L2 through interference filter 30 and is focused onto substrate 6 via a microscope objective lens L1. The radiation is then reflected via partially reflective surface 34 into a CCD detector in the back projected pupil plane 40 in order to have the scatter spectrum detected. The pupil plane 40 is at the focal length of the lens system L1. A detector and high NA lens are placed at the pupil plane. The pupil plane may be re-imaged with auxiliary optics since the pupil plane of a high NA lens is usually located inside the lens.

[0045] The pupil plane of the reflector light is imaged on the CCD detector with an integration time of, for example, 40 milliseconds per frame. In this way, a two-dimensional angular scatter spectrum of the substrate target is imaged on the detector. The detector may be, for example, an array of CCD detectors or CMOS detectors. The processing of the spectrum gives a symmetrical detection configuration and so sensors can be made rotationally symmetrical. This allows the use of a compact substrate table because a target on the substrate can be measured at any rotational orientation relative to the sensor. All the targets on the substrate can be measured by a combination of a translation and a rotation of the substrate.

[0046] A set of interference filters 30 may be available to select a wavelength of interest in the range of, say, 405 - 790 nm or even lower, such as 200 - 300 nm. The interference filter may be tunable rather than comprising a set of different filters. A grating could be used instead of one or more interference filters.

[0047] The substrate 6 (or even the reflective surface 34) may be a grating. The grating may be printed such that after development, a series of bars are formed of solid resist lines. The bars may alternatively be etched into the substrate. This pattern is sensitive to comatic aberrations in a lithographic projection apparatus, particularly the projection system PL, and illumination symmetry and the presence of such aberrations will manifest themselves in a variation in the printed grating. Accordingly, the scatterometry data of the printed gratings is used to reconstruct the gratings. One or more parameters of the grating, such as line widths and shapes, may be input to the reconstruction process from knowledge of the printing step and/or other scatterometry processes.

[0048] In transmission metallic gratings with rectangular slits, complex photonic band structures (CPBS) are shown to exhibit strong discontinuities, which are located on Wood-Rayleigh anomalies and reveal two types of resonance, which are referred to as horizontal and vertical surface-plasmon resonances. Spectral position and width of peaks in the spectrum can be directly extracted from CPBS for both horizontal and vertical resonances. In this way, the radiation coming off a transmission metallic grating can have its spectrum analyzed and one or more properties of the grating determined by the strong discontinuities located on the Wood-Rayleigh anomalies. Wood-Rayleigh anomalies occur upon the variation of wavelength or angle of incidence, giving an additional propagating diffraction order. The greater the beam width, the greater the lateral displacement of the beam.

[0049] An embodiment of the present invention detects the spectrum and creates a symmetrical pupil plane image from which the discontinuities can be measured and one or more grating properties therefore calculated.

[0050] According to an embodiment of the invention, the scatterometer may be adapted to measure the overlay of two misaligned periodic structures by measuring asymmetry in the reflected spectrum, the asymmetry being related to the extent of the overlay.

[0051] In an embodiment, the scatterometer is adapted to measure the overlay of two misaligned gratings or periodic structures by measuring asymmetry in the reflected spectrum and/or the detection configuration, the asymmetry being related to the extent of the overlay. Thanks to the symmetrical detection configuration, any asymmetry is clearly distinguishable. This provides a straightforward way to measure misalignment in the gratings.

[0052] One type of substrate pattern used is shown in Figure 4. A grating 14 has a second grating 12 printed on top of it. The amount by which the grating 12 is offset with respect to grating 14 is known as the overlay 22.

[0053] Note that in the embodiment shown in Figure 4a, the radiation source 2 illuminates the object symmetrically with respect to the surface normal and the scatterometry detector measures scatter radiation from several angles, although a source which illuminates the object from an oblique angle is also possible.

[0054] Overlay metrology is based on the measurement of an asymmetry in the angular scatter spectrum. Symmetric structures yield symmetric angular spectra and an asymmetry in the target shows up as an asymmetry in the angular scatter spectrum. This property is the basis of overlay metrology using angle-resolved scatterometry.

[0055] Two overlapping but misaligned gratings 12 and 14 made of bars with width 20 form one composite asymmetric target. The resulting asymmetry in the angular scatter spectrum is detected with the angle-resolved scatterometer 4 shown in Figure 3 and used to derive the overlay 22 in the following manner:

[0056] Two grating pairs are used with a deliberate bias of $+d$ and $-d$ in, respectively, the first and second pair. In other words, grating 12 is shifted in one direction in one pair (as shown in Fig. 4) and in the opposite direction in the other pair (not shown). The actual transverse shift between the gratings in each pair is therefore $X_1 = OV+d$ and $X_2 = OV-d$, OV being the overlay 22.

[0057] When the grating pairs are aligned, the overlay is 0 and if the intensity of the illumination incident on the gratings is I_{III} and the intensity of the radiation reflected off the gratings is I_{+1} in a first direction and I_{-1} in the opposite direction but in the same plane, when the overlay, OV = 0,

$$I_{+1} = I_{-1}. \qquad (1)$$

[0058] However, if

$OV \neq 0$,

$I_{+1} \neq I_{-1}$. (2)

[0059] For a small overlay, the intensity difference is proportional to the overlay:

$I_{+1} - I_{-1} = K \times OV$. (3)

K is a constant and is process dependent and therefore unknown.

[0060] In order to calibrate the overlay metrology with the scatterometer according to an embodiment of the present invention, two grating targets are used; one with the overlay shown in Figure 4b and a second with the exact reverse overlay, so the upper grating 12 is displaced to the left rather than the right with respect to the bottom grating 14. The overlay in the first set-up is OV+d (distance 22 in Figure 4b) and the overlay in the second set-up is OV-d.

[0061] So, for

$OV+d$,

asymmetry

$A_+ = K(OV+d)$ (4)

and for

$OV-d$,

asymmetry

$A_- = K(OV-d)$. (5)

[0062] The scaling factor K can be eliminated:

$$OV = d\frac{A_+ + A_-}{A_+ - A_-} \quad (6)$$

[0063] The overlay can therefore be calculated using measurements of the asymmetry in the angle resolved scatter spectrum.

[0064] An advantage of this method compared to previously known methods is the fact that only two gratings are required. Moreover, in principle, the method can also work for 2-D gratings: in that case only 2 gratings are required for a complete (x,y) overlay measurement. This is a significant improvement compared to, say, 6 gratings that spectroscopic scatterometry methods use.

[0065] The analysis of xy overlay metrology using 2-D gratings is as follows:

[0066] Two gratings have an amplitude transmission of $f(x,y)$ and $g(x,y)$. These gratings are periodic in two directions and their transmissions can therefore be written as a Fourier series:

$$f(x,y)=\sum_n\sum_m F_{n,m}\mathrm{e}^{-\mathrm{j}(nx+my)}$$
$$g(x,y)=\sum_p\sum_q G_{p,q}\mathrm{e}^{-\mathrm{j}(px+qy)} \tag{7}$$

[0067] Both gratings have an equal period and for simplicity the periods of the gratings have been normalized to 2π for the following calculations. The coefficients $F_{n,m}$ and $G_{p,q}$ can be interpreted as diffraction efficiencies that depend on the grating shape, wavelength and polarization. The two gratings overlap with a relative overlay of x_0 and y_0 in, respectively, the x and y directions. The total transmission t can be written as:

$$\begin{aligned} t(x,y) &= f(x,y)g(x-x_0,y-y_0) \\ &= \sum_n\sum_m\sum_p\sum_q F_{n,m}G'_{p,q}\mathrm{e}^{-\mathrm{j}((p+n)x+(q+m)y)} \end{aligned} \tag{8}$$

where:

$$G'_{p,q}=G_{p,q}\mathrm{e}^{\mathrm{j}(px_0+qy_0)} \tag{9}$$

[0068] The variables can be adjusted as follows:

$$p+n=a \Rightarrow p=a-n$$
$$q+m=b \Rightarrow q=b-m$$

[0069] Substituting these expressions in the Fourier series of $t(x,y)$ yields:

$$\begin{aligned} t(x,y) &= \sum_n\sum_m\sum_p\sum_q F_{n,m}G'_{n,m}\mathrm{e}^{-\mathrm{j}((p+n)x+(q+m)y)} \\ &= \sum_a\sum_b T_{a,b}\mathrm{e}^{-\mathrm{j}(ax+by)} \end{aligned} \tag{10}$$

where:

$$T_{a,b}=\sum_n\sum_m F_{n,m}G'_{a-n,b-m} \tag{11}$$

$T_{a,b}$ can be interpreted as the amplitude of the diffraction order (a,b). It can be see that this amplitude generally depends on the overlay in the x and y direction.

[0070] For simplicity, only diffraction orders running in the x-direction are considered. The analysis that follows can also be done for diffraction orders in the y-direction. This would only require an adjustment of variables.

[0071] For diffraction orders that run in the x-direction, b=0, so for the amplitude of two diffraction orders a and $-a$:

$$\begin{aligned} T_{a,0} &= \sum_n \sum_m F_{n,m} G_{a-n,-m} e^{j((a-n)x_0 - my_0)} \\ T_{-a,0} &= \sum_n \sum_m F_{n,m} G_{-a-n,-m} e^{j((-a-n)x_0 - my_0)} \end{aligned} \tag{12}$$

taking the factor $e^{\pm jax_0}$ in front of the summation yields:

$$\begin{aligned} T_{a,0} &= e^{jax_0} \sum_n \sum_m F_{n,m} G_{a-n,-m} e^{-j(nx_0 + my_0)} \\ T_{-a,0} &= e^{-jax_0} \sum_n \sum_m F_{n,m} G_{-a-n,-m} e^{-j(nx_0 + my_0)} \\ &= e^{-jax_0} \sum_{-n} \sum_m F_{-n,m} G_{-a+n,-m} e^{j(nx_0 - my_0)} \end{aligned} \tag{13}$$

[0072] Assuming that both gratings are symmetric in the x-direction:

$$\begin{aligned} F_{-n,m} &= F_{n,m} \\ G_{-n,m} &= G_{n,m} \end{aligned} \tag{14}$$

[0073] Using this property yields for the diffracted amplitudes:

$$\begin{aligned} T_{a,0} &= e^{jax_0} \sum_n \sum_m F_{n,m} G_{a-n,-m} e^{-j(nx_0 + my_0)} \\ T_{-a,0} &= e^{-jax_0} \sum_n \sum_m F_{n,m} G_{a-n,-m} e^{j(nx_0 - my_0)} \end{aligned} \tag{15}$$

[0074] The scatterometer measures the intensities of the diffracted fields, giving:

$$I_{\pm a,0} = \left| T_{\pm a,0} \right|^2 \tag{16}$$

[0075] Evaluation of this expression shows that the intensity can be written in the form:

$$\begin{aligned} I_{a,0} &= \sum_n \sum_m B_{n,m} \cos(\varepsilon_{n,m} - nx_0 - my_0) \\ I_{-a,0} &= \sum_n \sum_m B_{n,m} \cos(\varepsilon_{n,m} + nx_0 - my_0) \end{aligned} \tag{17}$$

where the amplitudes $B_{n,m}$ and phases $\varepsilon_{n,m}$ depend on the grating shapes, illumination wavelength and illumination polarization. Taking the difference of the +1 and -1 order yields an asymmetry A_x that runs in the x-direction:

$$\begin{aligned} A_x &= I_{1,0} - I_{-1,0} \\ &= \sum_n \sum_m B_{n,m} \cos(\varepsilon_{n,m} - nx_0 - my_0) - \sum_n \sum_m B_{n,m} \cos(\varepsilon_{n,m} + nx_0 - my_0) \\ &= \sum_n \sum_m 2B_{n,m} \sin(\varepsilon_{n,m} - my_0) \sin(nx_0) \end{aligned} \tag{18}$$

[0076] In practice the overlay is small compared to the pitch of the gratings. For example, the pitch is often of the order of 1 μm and the maximum overlay is of the order of 60 nm. The expression above can therefore be linearized and only the linear terms in x_0 and y_0 retained:

$$\begin{aligned} A_x &= \sum_n \sum_m 2B_{n,m} \sin(\varepsilon_{n,m} - my_0) \sin(nx_0) \\ &= \sum_n \sum_m 2B_{n,m} [\sin(\varepsilon_{n,m}) \cos(my_0) - \cos(\varepsilon_{n,m}) \sin(my_0)] \sin(nx_0) \\ &\cong \sum_n \sum_m 2B_{n,m} [\sin(\varepsilon_{n,m}) - \cos(\varepsilon_{n,m}) my_0] nx_0 \\ &= x_0 K_0 + K_{xy} x_0 y_0 \end{aligned} \tag{19}$$

where

$$\begin{aligned} K_0 &= \sum_n \sum_m 2nB_{n,m} \sin(\varepsilon_{n,m}) \\ K_{xy} &= \sum_n \sum_m 2mnB_{n,m} \cos(\varepsilon_{n,m}) \end{aligned} \tag{20}$$

[0077] It can be seen that there is a coupling term: The asymmetry in the x-direction is also a function of the y-overlay via the coupling term K_{xy}. If the 2-D grating has 90° rotation symmetry and if the light is polarized at 45°, then we can write for the asymmetry in the x and y directions:

$$\begin{aligned} A_x &= x_0 K_0 + K_{xy} x_0 y_0 \\ A_y &= y_0 K_0 + K_{xy} x_0 y_0 \end{aligned} \tag{21}$$

[0078] These equations are the basis for xy overlay metrology with two 2-D grating pairs. In the first grating pair, a bias of $+d$ is introduced in the upper grating and in the second grating pair, a bias of $-d$ is introduced. This bias is applied in both the x and y

direction. Four asymmetry terms can now be measured: An x and y asymmetry in the first grating pair and an x and y asymmetry in the second grating pair are shown as:

$$\begin{aligned} A_{1x} &= K_0(OV_x + d) + K_{xy}(OV_y + d)(OV_x + d) \\ A_{1y} &= K_0(OV_y + d) + K_{xy}(OV_y + d)(OV_x + d) \\ A_{2x} &= K_0(OV_x - d) + K_{xy}(OV_y - d)(OV_x + d) \\ A_{2y} &= K_0(OV_y - d) + K_{xy}(OV_y - d)(OV_x - d) \end{aligned} \tag{22}$$

[0079] This gives four non-linear equations with four unknowns K_0, K_{xy}, OV_x and OV_y which can be solved to give the overlay.

[0080] In an embodiment, one or more apertures may be provided to the scatterometer to mimic lithography exposure conditions when the grating pattern(s) was created. The apertures may then be used in the creation of the angle-resolved spectroscopic image of the grating pattern(s) using the scatterometer.

[0081] In an embodiment, it is possible to immerse at least part of the space between the substrate and the detector in liquid, more specifically, the space between lens L1 and the substrate 6 as shown in Figure 3. The liquid may be water. This has the advantage of increasing the spatial bandwidth of the medium between the substrate 6 and the lens L1. This means that a diffraction that would be evanescent, for example, in air can propagate and be captured by the lens. With immersion of the space, therefore, it becomes possible to detect a higher diffraction order that contains more detailed information about the grating under investigation than with, for example, air in the space.

[0082] The numerical aperture (NA) of the scatterometer is preferably at least 0.9, even 0.95 or above 1.

[0083] Immersing the space between L1 and the object with a high refractive index fluid increases the spatial bandwidth of the medium and allows the propagation of a higher diffraction order for smaller pitches. The smallest pitch that creates a propagating first order spectrum is $\frac{\lambda}{(2NA)}$. Assuming NA equals 1.3 and λ equals 400 nm, this yields a minimum pitch of 154 nm. This corresponds to a critical dimension (CD) or reconstructed grating width of approximately 20 to 80 nm. When looking at a profile such as that shown in Figure 2, the critical dimension is the mean width of a peak and the pitch is the distance from one peak to the next.

[0084] The immersion fluid should have a large index step with respect to, for example, the resist that is on the substrate 6. This may allow maximum contrast in the detector image. A possible liquid that fulfils such requirements is water.

[0085] Figure 5 shows, according to an embodiment of the invention, the use of one and the same detector to monitor the source output intensity and the intensity of scattered radiation, which avoids synchronization problems and allows a real-time compensation for source output variations.

[0086] The scatterometer may comprise a non-polarizing beam splitter and a tilted mirror for coupling off a portion of the radiation beam emitted from the radiation source for separate measurement with the same detector. In an embodiment, the portion of the radiation beam is used to measure the intensity of the radiation beam and the scatterometer may be adapted to compensate for fluctuations in intensity of the radiation beam. Advantages of using the same CCD detector for the intensity measurement beam alongside the main measurement beam are that no extra detector is required and so there is no difference in optical and thermal properties between a reference sensor and a metrology sensor; and there are no extra electronics required to trigger, read out and store the reference signal. Any intensity variations may be measured and compensated for.

[0087] A non-polarizing beam splitter 50 in the radiation path images scattered radiation on a two-dimensional detector 32. An extra lens re-images the pupil plane onto the CCD detector. The intensity incident on the detector is shown as image 36. The non-polarizing beam splitter 50 also couples out a portion of the radiation beam to use it for monitoring intensity noise. Instead of measuring this radiation portion with a separate detector, the light is retro-reflected using tilted mirror 52 and transmitted to a separate part of the same detector 32. An optional pupil stop 54 limits the extent of the radiation portion and the mirror tilt ensures that the radiation portion is projected alongside the main radiation beam. The spectrum is imaged onto the detector 32 at the pupil plane of L1.

[0088] In previous methods, angle-resolved scatterometry has been done at a single wavelength. Measurements at different wavelengths would then have been done sequentially and the different wavelengths would be time multiplexed. However, time multiplexing of the wavelengths may degrade throughput.

[0089] In an embodiment, the scatterometer comprises a wavelength multiplexer between the radiation source and the substrate and a demultiplexer between the substrate

and the detector. This allows several different wavelengths (or colors) to be measured simultaneously, giving more information in a shorter time frame and therefore robustness as discussed above. The wavelength multiplexer may comprise a dispersive element placed at a back-projected object plane or a dispersive element placed at a pupil plane.

[0090] The surface area of the radiation source is preferably split into N parts that are each coupled to a wavelength multiplexer, where N is the number of discrete wavelengths. This splitting can be done, for example, with fiber bundles and the like.

[0091] In an embodiment, the multiplexer comprises a dispersive element placed at a back-projected object plane. The dispersive element may be a grating or prism adapted to accommodate N discrete wavelengths each with a bandwidth $\delta\lambda$ and a spacing of at least twice the bandwidth, *i.e.* 2 $\delta\lambda$. This may maximize the usage of an extended light source. Measurement of different wavelengths no longer has to be time-multiplexed because it can be done at the same time, and so a major advantage is that throughput is increased.

[0092] In an embodiment, the demultiplexer comprises a dispersive element placed at a pupil plane. One or more optical wedges may be inserted in the object plane to achieve well-defined separation of angle-resolved spectra in the pupil plane.

[0093] In an embodiment, an extended broadband radiation source such as a xenon, deuterium or quartz tungsten halogen light source is used. These sources have a large etendue that gives a surface area that can be split into discrete wavelengths and offer more information as discussed above. The wavelengths may be in the range of 193 to 800 nm.

[0094] In an embodiment, a dispersive prism or grating which combines N discrete wavelengths is used in the illumination branch (or the radiation path between the source 2 and the substrate 6 in Figure 2) and a grating or prism is used in the detection branch (or the space between the radiation path between the substrate 6 and the detector 4) to spatially separate the wavelengths.

[0095] An example of a multiplexing grating is shown in Figure 6. Two light sources S1 and S2 are transmitted through a lens system L2 and strike a Littrow mounted grating 16 which is in the object plane 42 and are focused on the pupil plane 40 before being transmitted through a lens system L1 to another object plane 42 and optionally into an illumination fiber 60. The pupil plane contains rectangular apertures of suitable dimensions - the width determining the angular extent of the light incident on the grating. This angular extent and the grating pitch determine the bandwidth of the returning light that

is transmitted via the aperture in the pupil plane. For example, a grating with 1200 lines per millimeter yields a dispersion of approximately 1.2 mrad/nm. An effective bandwidth of 4 nm corresponds to a full angular width of the illumination beam of 3.33 mrad. The spot size of the substrate 6 is approximately 40 μm and its NA is 0.95. The beam diameter on the grating is therefore approximately 23 mm. If the focal length of the lens L1 is 100 mm, then the width of the aperture holes in the pupil plane must be 333μm. If an illumination fiber is used, then the illumination NA must be approximately 0.22.

[0096] Clearly more than two radiation sources (with different wavelengths) may be used at a time.

[0097] Figure 7 shows an example of a wavelength demultiplexer in the detection branch. For simplicity, the separation of only two wavelengths is again shown. The demultiplexer is similar to the muliplexer, except that the grating is placed in the pupil plane and not in the object plane. The light that is diffracted by the grating in the Littrow mounted grating 16 is transmitted by the lens L2, which makes two object images with wavelengths $\lambda 1$ and $\lambda 2$ in the object plane. This plane may contain field stops with n holes (n = 2 in this case), which should be sufficiently wide to avoid spatial filtering to avoid disturbing the spectrum. Each hole of the field stop 40 also has a wedge 62 with a unique wedge angle. This wedge 62 ensures that the angle resolved scatter spectrum for each wavelength is imaged on a different part of the CCD detector 32. The CCD detector is based at the second pupil plane 40.

[0098] Since the wedges 62 can deflect the light in two directions, it is possible to realize an efficient filling of a CCD detector with many angle resolved spectra.

[0099] In order to obtain reproducible results, the targets should be well focused. In order to achieve this, the pupil plane 40 of a high NA objective is imaged on a detector with a double telecentric system as shown in Figure 8 according to an embodiment of the invention. In all embodiments, the NA is high; preferably at least 0.9.

[00100] A knife edge 70 in the intermediate object plane 42 blocks one half of the intermediate object image. The edge may be a Foucault knife-edge.

[00101] The knife-edge helps to focus the image of the radiation onto the substrate. For each orientation, the intensity in the outer regions (or practically speaking, in two halves) of the pupil plane 40 is sampled. In the case of a defocus, a difference in intensities I1 and I2 is produced. Focus F is given as:

$$F = k\frac{I1 - I2}{I1 + I2} \qquad (23)$$

[00102] The proportionality factor k is independent of the image and needs to be determined only once, though since the focus sensor can be used in an integrating feedback loop, the precise value of k is not important.

[00103] Illumination sources are not always homogeneous and must be calibrated and corrected for in order to guarantee precise exposure of the substrates. Inhomogeneity may be caused by the illumination sources themselves, or by the roughness of surface coatings of reflectors in the illumination path. Previously, normalization of the illumination beam has been done using an aluminium mirror. However, this normalization fails when the object to be measured (i.e. gratings or the substrate) generate higher diffraction orders. These cause tool induced shift errors in overlay applications.

In an embodiment, therefore, the scatterometry system further comprises one or more mirrors in the illumination beam. More specifically, the mirrors may be a fiducial on the substrate table and may be made of aluminium. These mirrors either tilt or exist at different tilt angles in order to create at least two images reflected in different angles. For each tilt angle, the detected spectrum shifts in the same direction as the direction of the tilt. These images are detected and combined into differential equation, from which the illumination profile of the radiation beam may be determined. The resultant illumination profile is used to correct measurements of the property of the reflected spectrum at higher diffraction orders.

[00104] The measured signal, $M_0(k)$, is represented as:

$$M_0(k) = [A(-k)R_0(k) + A(-k \pm k_G)R_{\mu 1}(x) + \Lambda + A(-k \pm Nk_G)R_{\mu N}(x)]B(k) \qquad (24)$$

[00105] where:

$A(k)$ is the unknown illumination intensity at position k in the pupil plane;

$B(k)$ is the unknown optical loss in the detection branch of the sensor; and

$R_{\pm N}$ is the diffraction efficiency of the Nth order of the grating object.

[00106] In practice, the illumination intensity varies because of a slowly varying inhomogeneous illumination beam and surface roughness of the optics and coatings in the illumination path. The surface roughness of optical coatings generally gives rise to a grainy

appearance of the illumination beam in the pupil plane.

[00107] A reference measurement may be carried out on a highly reflecting aluminum mirror, which yields the following measured signal:

$$M_M(k) = A(-k)R_M(k)B(k) \tag{25}$$

[00108] Normalizing the measurement of the object with the reference yields:

$$\frac{M_0(k)}{M_M(k)} = \frac{R_0(k)}{R_M(k)} + \frac{A(-k \pm k_G)}{A(-k)}\frac{R_{\mu 1}(x)}{R_M(k)} + \Lambda + \frac{A(-k \pm Nk_G)}{A(-k)}\frac{R_{\mu N}(x)}{R_M(k)} \tag{26}$$

[00109] It can be seen that the losses in the detection branch are eliminated by this normalization.

[00110] However, inhomogeneities in the illumination are only eliminated for the zero diffraction order (i.e. the specular reflection). Higher diffraction orders retain an unknown error term of the form:

$$\frac{A(-k \pm Nk_G)}{A(-k)} \tag{27}$$

[00111] In order to eliminate this term, the illumination profile $A(k)$ must be calibrated as discussed below.

[00112] The mirror may be a single convex or concave mirror, or it may be a plane mirror that is actively tilted over a range of angles during detection. Alternatively, there may be a range of mirrors at different tilt angles. The measured reflection angle may be in a radial direction (this alters the magnitude of the tilt) or in an azimuthal direction (this alters the direction of the tilt).

[00113] The method used to determine the differential equation will now be described in 1-dimension. The extension to 2 dimensions is trivial.

[00114] A reference mirror is measured for two small opposite mirror tilts $\pm\theta_M$ of the order of 1 mrad. As a result of this tilt, the measured pupil image will shift. Two slightly shifted images are therefore measured:

$$M_{\pm\theta}(k) = A(-k \pm \Delta(k))R_M(k)B(k)C(k;\pm\theta) \tag{28}$$

[00115] Here, Δ is the shift in the pupil plane, which generally depends on the

position k in the pupil plane. For an aplanatic system:

$$\Delta(k) = 2\theta_M \sqrt{1-k^2} \tag{29}$$

[00116] C in equation (18) accounts for the redistribution of the reflected intensity and for an aplanatic system:

$$C(k;\Delta) = 1 + \frac{2\theta_M k}{\sqrt{1-k^2}} \tag{30}$$

$$Q_M(k) = \frac{M_{+\theta} - M_{-\theta}}{M_{+\theta} + M_{-\theta}} \tag{31}$$

where $M_{+\theta}$ and $M_{-\theta}$ are spectra measured at a small positive tilt and small negative tilt respectively.

[00117] Here, the subscript 'M' of Q is used to emphasize that it concerns measured data. For small tilts, an approximation may be:

$$A(k + \Delta(k)) \cong A(k) + \frac{\mathrm{d}A}{\mathrm{d}k}\Delta(k) \tag{32}$$

[00118] Using this linearization yields for Q the differential equation:

$$\frac{Q(k)}{\Delta(k)} = \frac{1}{A(k)}\frac{\mathrm{d}A}{\mathrm{d}k} \tag{33}$$

[00119] This equation is easily solved to yield:

$$A(k) = \exp\left[\int_0^k \frac{Q(k')}{\Delta(k')}\mathrm{d}k'\right] \tag{34}$$

[00120] The above derivation can be easily extended to 2 dimensions. In practice, the measured data is not continuous but is digitized sampled data. However, this does not alter the concept derived above.

[00121] In practice, a plane mirror may be employed that is mechanically tilted using actuators. A more elegant and simple approach is the use of a concave or convex mirror with a radius of curvature R and lateral position x. The local height of a curved mirror z is described by:

$$z = \frac{x^2}{2R} \tag{35}$$

[00122] The local slope of the surface θ scales linearly with the lateral position x:

$$\begin{aligned} \theta &\cong \frac{\mathrm{d}z}{\mathrm{d}x} \\ &= \frac{x}{R} \end{aligned} \tag{36}$$

A concave or convex spherical aluminum fiducial on the substrate stage therefore renders the calibration straightforward because the proper tilt is simply achieved by moving the fiducial to the proper location under the detector. An embodiment of the invention uses a radiation beam with an annular intensity distribution in a conjugate plane to the substrate. In order to create the annular intensity distribution, the radiation source may comprise mechanical blades, spatial light modulators or spatially coherent broadband lasers and a zoom-axicon (i.e. to create a ring of laser light). the annular radiation beam preferably comprises small-Φ illumination.

[00123] Implementing annular radiation has advantages over, say, inserting a blade, because there is no light loss because almost all the photons are "used". This is particularly important where light sources such as UV or DUV are used because they emit fewer photons than more abundant light sources and so losing a number of those photons is more noticeable. In particular, this is noticeable in signal collection because the lithographic tool suffers a certain amount of delay if there is a lower light intensity. Annular light sources have the further advantage of not causing internal reflections as blades might. Internal reflections require blocking to avoid light artefacts. Of course, other illumination techniques, such as quadrupole illumination, which offer the same advantages may be used.

[00124] Ideally, the annulus of the annular radiation is placed in the pupil plane of the high NA lens. However, the pupil plane is not directly accessible and in practice, the annulus is placed in a back-projected image of the pupil plane in the illumination branch of the scatterometer. The advantage of annular illumination is that the intensity of the +1/-1 diffraction order of a grating with a small pitch of the order of λ/NA may be separately measured.

[00125] This embodiment may be used for calculating variations in substrate tilt by putting a shaped obscuration in the radiation beam and detecting changes in the width and

shape of the shaped obscuration on the substrate caused by variations in the substrate tilt. The shaped obscuration may be, for example, a cross-hair as shown in Figures 9a and 9b. It may, of course, be any other shape and it does not have to be in the centre of the pupil plane.

[00126] The idea of measuring wafer tilt is based on the fundamental relation that a tilt in the wafer plane causes a shift in the pupil plane. In the present embodiment, a cross-haired obscuration is placed in the center of the illumination beam. This produces a black cross-hair in the scattered light in the pupil plane as shown in Figure 9a.

[00127] The location of this cross will vary if the substrate tilt changes. As a result, the difference may be measured between this pattern (at zero tilt) and an actual measurement at an unknown tilt to obtain an image as shown in Figure 9b. A small tilt in the substrate does not lead to a substantial shape change in the annulus of light, but rather, it will lead to a shift of the pupil plane image. This shift is generally small and of the order of 0.1 pixels. In order to be able to detect such a small shift, the values between pixels may be interpolated by curve fitting, for example. By fitting a curve through the dark-light transition that occurs at the edge of the annulus, sub-pixel displacements of the annulus may be measured.

[00128] The width and sign of these transitions can be used to calculate and correct for the substrate tilt in 2 dimensions. In this way the substrate can be measured at constant (zero) tilt.

[00129] Figure 10 shows the diffraction orders of small pitched gratings detected using radiation with an annular intensity distribution in a conjugate plane to the substrate. Using an annular intensity distribution allows the shape of the images as shown in Figure 10 and thereby allows clearer and more precise measurement of substrate tilt. The image labelled 0 is the central zero-order diffraction order as imaged in the detector. The images labelled –2, -1, 1 and 2 are higher diffraction orders. These higher diffraction orders are shifted with respect to the lower diffraction order and so are easier to measure for overlay metrology of isolated 1-D and 2-D features.

[00130] In order to speed up calculation times, there are cases in which it is not necessary to calculate a simulated signal in every single position in the pupil plane, especially when smooth variations are expected. In these cases, a coarse grid may be measured and a pixel interpolation technique used to interpolate the overall shape at the

pupil plane. An annular beam is more advantageous in this case, too, because there are areas in the pupil plane that only receive light from first order diffraction. If a block beam were used, for instance, each point in the pupil plane would receive light from either the zeroth order or a combination of the zeroth order and the first order, causing errors in the measurement at the pupil plane.

[00131] Normal measurements using a scatterometer involve measuring the properties of a single target on a single substrate with a single polarization at one time. This limits the throughput of substrates through the scatterometry; and potentially exposure steps. An embodiment of the invention uses illumination source to project a plurality of illumination spots onto a substrate. The detector of the scatterometer simultaneously detects an angle-resolved spectrum of the plurality of illumination spots reflected from a surface of the substrate. The plurality of illumination spots may be created using a double illumination fiber or a Wollaston prism for creating two orthogonally polarized illumination spots.

[00132] Figure 11 shows part of the scatterometer hardware. Two illumination spots 70 are split in beam splitter 50 before being transmitted down through the high numerical aperture objective positioned in the pupil plane 40 onto the substrate 6. The reflected beam is transmitted upwards through two wedges 62 that separate the two angle resolved spectra in the pupil plane, the wedges themselves being positioned in the intermediate image plane 42. The illumination beams are then detected by the CCD on the re-imaged pupil plane 40 at the top of Figure 11. Two, or even more, parallel measurements may thereby be made; for example, of horizontal and vertical lines for a single polarization; or even for both horizontal and vertical lines for both TE and TM polarization.

[00133] An embodiment of the invention converts the scatterometer into an ellipsometer. In order to do this, the illumination branch further includes a first polarizer configured to linearly polarize the radiation beam; a beam splitter configured to split the radiation beam into two orthogonal components (E_{TE}, E_{TH}); a second polarizer configured to polarize the scattered beam; a variable compensator (a Pockells Cell, Wollaston prism pair or Soleil-Babinet compensator) positioned between the first and second polarizers, the variable compensator being configured to vary the optical path difference between orthogonally polarized components (and optionally between the beam splitter and the high numerical aperture lens); and a 2-dimensional detector for detecting sinusoidal intensity

variation of the resultant beam components. The compensator is most commonly in the main illumination branch of the scatterometer, though it may of course be in a second illumination branch.

[00134] The 2-dimensional detector such as a Complementary Metal Oxide Semiconductor detector (CMOS) has a high frame rate, i.e. in the region of 1000 frames per second.

[00135] Figure 12 shows how the angular-resolved spectroscopic concept is turned into an angle-resolved spectroscopic ellipsometer. An illumination beam with two wavelengths, λ_1 and λ_2 is transmitted through a 45° polarizer 72, reflected off the substrate 6 and transmitted through a variable compensator 74 before being polarized again. Between the beam splitter and the variable compensator 74, the illumination beam is divided into two beams with a phase difference Δ between the TE and TM polarized beams. The grid 36 in Figure 12 shows the 2-D detector array and the time-dependent intensity variation in one pixel of this array. The other pixels will show a comparable variation. The beams are passed through two bandpass filters 76 to obtain the illumination profiles of λ_1 and λ_2. The resultant ellipsometric parameters $\cos(\Delta)$, $\sin(\Delta)$ and $\tan(\Psi)$ are virtually insensitive to internal sensor scattering and so the signal to noise ratio can be improved. The operation is modeled with Jones vectors and matrices below, though it could also be modeled using Mueller matrices, which enable the inclusion of imperfections of the optical components in the mathematical models.

[00136] The illumination field after the first polarizer is 45 ° polarized and described by the Jones vector:

$$E_{inc} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{37}$$

[00137] Basis vectors correspond to TE and TM polarized light that is incident on a target on a sample. The act of reflecting off the sample causes an alteration in the amplitude and the phase of the scattered TE and TM components. This can be represented by a Jones matrix:

$$J_{sample} = \begin{bmatrix} R_{TE} & 0 \\ 0 & R_{TM} e^{j\Delta} \end{bmatrix} \tag{38}$$

where Δ is the phase difference between the TE and TM components of the scattered fields and R_{TE} and R_{TM} are the amplitudes of, respectively, the scattered TE and TM fields.

[00138] These parameters are a function of angle of incidence and wavelength. Ignoring any phase and amplitude variations introduced by the high-NA lens and the beam splitter, for the incident field on the compensator:

$$E_{c_in} = J_{sample} E_{inc} = \begin{bmatrix} R_{TE} \\ R_{TM} e^{j\Delta} \end{bmatrix} \tag{39}$$

[00139] The compensator introduces a time-varying optical path difference (OPD)-variation between the TE and TM components. If the wavelength of the light is λ, for the Jones matrix of the compensator:

$$J_{comp} = \begin{bmatrix} 1 & 0 \\ 0 & e^{j2\pi \frac{OPD(t)}{\lambda}} \end{bmatrix} \tag{40}$$

and so the field after the compensator is:

$$E_{c_out} = J_{comp} E_{c_in} = \begin{bmatrix} R_{TE} \\ R_{TM} e^{j\left(\Delta + 2\pi \frac{OPD(t)}{\lambda}\right)} \end{bmatrix} \tag{41}$$

[00140] The polarizer is oriented at 45° and has a Jones matrix:

$$J_{pol} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{42}$$

and so the field after the polarizer is:

$$E_{pol_out} = J_{pol} E_{c_out} = \frac{1}{2}\begin{bmatrix} R_{TE} + R_{TM} e^{j\left(\Delta + 2\pi \frac{OPD(t)}{\lambda}\right)} \\ R_{TE} + R_{TM} e^{j\left(\Delta + 2\pi \frac{OPD(t)}{\lambda}\right)} \end{bmatrix} \tag{43}$$

[00141] The intensity incident on the detector array is:

$$I_d = E^T_{pol_out} \cdot E^*_{pol_out} = \frac{1}{2}\left[R_{TE}^2 + R_{TM}^2 + 2R_{TE}R_{TM}\cos\left(\Delta + 2\pi \frac{OPD(t)}{\lambda}\right)\right] \tag{44}$$

[00142] If the OPD increases linearly over the measurement time interval,

$OPD = Kt$

[00143] This yields a time-harmonic intensity variation:

$$I_d = \frac{1}{2}\left[R_{TE}^2 + R_{TM}^2 + 2R_{TE}R_{TM}\cos(\Delta + \Omega t)\right] \tag{45}$$

where

$$\Omega = 2\pi \frac{K}{\lambda} \tag{46}$$

[00144] The contrast of the intensity variation is directly related to the ellipsometric parameter tan(Ψ) and the phase of the sinusoidal variation directly yields the ellipsometric parameters cos(Δ) and sin(Δ). In a standard ellipsometric scatterometer, tan(Ψ) and cos(Δ) are the signals that are measured and simulated to obtain the profile information. In that case, tan(Ψ) and cos(Δ) are recorded as a function of wavelength. In the present invention, tan(Ψ) and cos(Δ) are obtained as a function of position in the pupil plane and can be used for similar analyses. In particular, the ellipsometric parameters are used to measure layer thickness by solving an inverse scattering problem, i.e. the measured parameters are compared with modeled parameters and the stack parameters are determined by minimizing the root-mean-square difference (or any other suitable metric) between the measured and modeled parameters.

[00145] Because the frequency of the variation depends on the wavelength, the various wavelengths can be separated with a bandpass filter. This can be easily realized via signal processing with, for example, discrete Fourier Transform techniques.

[00146] The compensator can also be place in the illumination branch. Moreover, it can also be placed between the beam splitter and a high numerical aperture objective. This has the advantage that the OPD variation is doubled.

[00147] The 2-D detector is the critical part of this concept; to ensure sufficiently short measurement times, it must have a high frame rate. CMOS detectors can achieve very high frame rates, for example 1000 frames per second.

[00148] Measuring overlay as described in paragraphs 51 to 80 above does not allow for the measurement of large overlay errors, in particular, overlay errors equal to an integer times the grating pitch. Clearly, the ability to detect overlay errors smaller than the grating pitch is of no use if there is an overlay error that means the grating lines are lined up with each other, but shifted by several grating pitch widths.

[00149] An embodiment of the invention therefore uses the second detector branch already present in the scatterometer (and discussed above) for carrying out coarse overlay measurements to determine whether coarse errors exist, such as whether the grating overlay error is in fact greater than the pitch of the grating. A coarse overlay measurement is an

imaging-based technique, wherein a second camera looks at an image of two overlapping gratings and determines whether there are large displacements by comparing the positions of the edges of markers on a substrate. A perfect overlay will have perfectly aligned marker edges. Pattern recognition algorithms are used to determine the edge of a grating in the process layer and the edge of the grating in the resist layer. This measurement is done on the four sides or corners of a grating pair. The measured edge positions are used to calculate the position of the resist grating relative to the position of the grating in the process layer.

[00150] The fact that scatterometry on its own cannot measure overlay that is equal to a multiple number of the grating pitch is a fundamental limitation because the measurement principle is based on grating coupling that varies periodically with the grating pitch. In other words, zero overlay and an overlay equal to the pitch yield identical results.

[00151] The scatterometer of the present invention allows a very simple solution. The scatterometer comprises a separate imaging branch that images the substrate surface on a CCD camera. This second camera branch is needed for measuring the position of the substrate through an alignment and pattern recognition step. The second branch is shown schematically in Figure 13.

[00152] The pupil plane 40 measurement (the actual angle resolved measurement) requires an illumination source that underfills the target at the object plane 42 (i.e. the measurement spot is smaller than the measurement target). The pupil plane imaging illumination beams are shown as solid lines in Figure 13. In this case, only a portion of the target is measured and structures outside the target area are not illuminated. If the measurement spot fills or overfills the measurement target, the measurement signal is disturbed by the area surrounding the target and data interpretation and signal analysis are unnecessarily complicated.

[00153] The image plane measurement, on the other hand, must overfill the target in order to detect the alignment because the complete pupil plane must be sampled, including the contours of the target. The light rays for the image plane measurements are shown as dashed lines. The image of the object plane is formed on a first CCD camera 80 and the image of the pupil plane is formed on a second CCD camera 82.

[00154] Figure 14 shows one possible example of an overlay target for zero overlay (left-hand drawing) and an X-overlay equal to twice the grating pitch (right-hand drawing).

The pupil plane measurement would yield the same measured overlay of zero for both situations making it an unreliable measurement. The image plane measurement, however, can clearly distinguish between these two situations. In this way, a robust two-stage metrology scheme may be carried out as follows:

[00155] (1) Two image plane measurements are carried out to verify that there are no large overlay errors present.

[00156] (2) If the previous measurement indicates that overlay is less than approximately 200 nm, the pupil plane measurements are carried out.

The 200 nm criterion is an indicative example. It may be made to any sensible threshold. Assuming that the image plane CCD has 1000 × 1000 pixels and assuming a pixel pitch of 100 nm at substrate level, the total field of view will be 100 × 100 μm^2, which is adequate for pattern recognition and alignment while still allowing coarse overlay measurements with an accuracy of the order of 20-50 nm.

[00157] Coarse overlay can only be measured when the entire alignment marker is visible to the CCD. If, for example, only the centre part of the marker is visible, the substrate table needs to be moved to the edge of the marker to enable the measurement of the coarse overlay. This calls for additional movement of the table, thereby slowing the measurement taking process. A larger field of view allows the capture of the marker in one "stroke" and a coarse measurement to be carried out quickly while a second camera is free to capture the image on the pupil plane and obtain the detailed overlay information.

[00158] The field of view that is required for capturing the relevant image can be reduced even further if the results of edge pre-alignment and coarse wafer align in the exposure tool are used. With these data, it is possible to predict the location of the markers on the substrate with μm accuracy after the edge pre-alignment in the overlay metrology module is complete.

[00159] An embodiment of the invention detects not just overlay errors, but also damaged gratings using the scatterometer arranged for CD metrology on gratings or other periodic structures. The scatterometer normally detects specular light; i.e. lowest order light that has been reflected directly off the grating. Local distortions in the grating destroy the periodicity of the grating and result in scattering in a non-specular direction. The scatterometer can be used to detect an angle-resolved spectrum of the scattered beam at various angles outside its specular direction. Radiation with an annular intensity

distribution, or small-Φ illumination, may be used for greater accuracy and images that easier to read.

[00160] An embodiment of the invention may be used for detecting bubble defects in an immersion lithographic apparatus, where a liquid is introduced between the projection system and the substrate as discussed above. Previously, bubble defects have been measured using off-line defect inspection tools. Off-line tools take a longer time to produce results than on-line tools because substrates must be taken out of the production line and queued. Bubbles in the liquid cause a surface imperfection on the substrate, which will cause light scattering when the surface is exposed to light. This scattered radiation is measured using the scatterometer of the invention and the cause of the scattering extrapolated back to bubble defects.

[00161] While specific embodiments of the invention have been described above, it will be appreciated that the invention may be practiced otherwise than as described. The description is not intended to limit the invention. The specifically described embodiments are extensions to a general operating principle and are not necessarily mutually exclusive; they are all combinable in a single metrology tool to increase its effectiveness based on results seen at a detector as described above. Further, although the embodiments described herein relate to lithography applications, the hardware and applications are not limited to these. They may be used for other applications such as monitoring etch process steps and the like.

CLAIMS

1. A scatterometer, comprising:

a lens configured to direct a radiation beam from a radiation source toward a pattern formed on a substrate; and

a detector located in a pupil plane of the lens and configured to detect an angle-resolved spectrum of the radiation beam reflected from the pattern,

wherein a property of the substrate is measured using the detector by measuring asymmetries between intensities of corresponding diffraction orders diffracted from the pattern in the reflected angle-resolved spectrum.

2. The scatterometer according to claim 1, wherein a numerical aperture of the lens is at least 0.9.

3. The scatterometer according to claim 1, wherein a numerical aperture of the lens is at least 0.95.

4. The scatterometer according to claim 1, wherein the detector is configured to measure at least one of (a) an intensity of a transverse magnetic and a transverse electric polarized light and (b) a phase difference between the transverse magnetic and the transverse electric polarized light.

5. The scatterometer according to claim 1, wherein asymmetries between intensities of corresponding diffraction orders diffracted from the pattern are a property of the reflected angle-resolved spectrum, wherein the property of the substrate is further measured by measuring, in the pupil plane of the lens, another property of the reflected angle-resolved spectrum at a plurality of wavelengths substantially simultaneously.

6. The scatterometer according to claim 5, wherein the plurality of wavelengths each has a spacing of at least twice a bandwidth of the plurality of wavelengths.

7. The scatterometer according to claim 1, further comprising an extended broadband radiation source configured to provide the radiation beam to the surface of the substrate, the radiation beam having a wavelength of at least 50 nm.

8. The scatterometer according to claim 1, wherein the asymmetries are related to an extent of an overlay between two misaligned periodic structures.

9. The scatterometer of claim 1, further comprising a non-polarizing beam splitter and a tilted mirror configured to couple off a portion of the radiation beam emitted from a radiation source for a separate measurement with the detector.

10. A scatterometer, comprising:

a high numerical aperture lens configured to direct a radiation beam from a radiation source toward a substrate;

a detector located in a pupil plane of the high numerical aperture lens and configured to detect an angle-resolved spectrum of the radiation beam reflected a surface of the substrate; and

a wavelength multiplexer located between a radiation source configured to provide the radiation beam and the substrate and a wavelength demultiplexer located between the substrate and the detector,

wherein a property of the substrate is measured using the detector by measuring an asymmetry in the reflected angle-resolved spectrum at a plurality of angles substantially simultaneously.

11. The scatterometer according to claim 10, wherein the wavelength multiplexer comprises a dispersive element placed at a back-projected object plane.

12. The scatterometer according to claim 11, wherein the dispersive element is one of a grating and a prism adapted to accommodate N discrete wavelengths each having a spacing of at least twice a bandwidth of the plurality of wavelengths.

13. The scatterometer according to claim 10, wherein a surface area of the radiation source is split into N parts that are each coupled to the wavelength multiplexer, where N is a number of discrete wavelengths.

14. The scatterometer according to claim 10, wherein the wavelength demultiplexer comprises a dispersive element placed at the pupil plane.

15. The scatterometer according to claim 1, further comprising an optical wedge in an object plane configured to achieve a pre-defined separation of the reflected angle-resolved spectrum in the pupil plane.

16. The scatterometer according to claim 1, wherein a portion of the radiation beam is used to measure an intensity of the radiation beam.

17. The scatterometer according to claim 1, the detector is adapted to compensate for fluctuations in an intensity pattern of the reflected radiation beam.

18. The scatterometer according to claim 1, further comprising a pupil stop configured to limit a size of a portion of the radiation beam.

19. A scatterometer, comprising:

a lens configured to direct a radiation beam from a radiation source toward a pattern formed on a substrate; and

a space between the substrate and the lens comprising a liquid,

wherein a property of the substrate is measured using a detector by measuring asymmetries between intensities of corresponding diffraction orders in an angle-resolved spectrum of a radiation beam diffracted from the pattern, at a plurality of angles and a plurality of wavelengths substantially simultaneously.

20. The scatterometer according to claim 19, wherein the liquid is water.

21. The scatterometer according to claim 19, wherein a numerical aperture of the lens is at least 0.9.

22. The scatterometer according to claim 19, wherein a numerical aperture of the lens is at least 0.95.

23. The scatterometer according to claim 19, wherein the detector is configured to measure at least one of (a) an intensity of a transverse magnetic and a transverse electric polarized light and (b) a phase difference between the transverse magnetic and the transverse electric polarized light.

24. The scatterometer of claim 19, further comprising a non-polarizing beam splitter and a tilted minor configured to couple off a portion of the radiation beam emitted from a radiation source for a separate measurement with the detector.

25. A scatterometer, comprising:

a high numerical aperture lens configured to direct a radiation beam from a radiation source toward a substrate;

a space between the substrate and the high numerical aperture lens comprising a liquid;

a wavelength multiplexer located between a radiation source, configured to provide the radiation beam, and the substrate; and

a wavelength demultiplexer located between the substrate and the detector located in a pupil plane of the high numerical aperture lens, and configured to measure another property of the reflected spectrum,

wherein a property of the substrate is measured using a detector by measuring an asymmetry in an angle-resolved spectrum of a radiation beam reflected from a surface of the substrate, at a plurality of angles and a plurality of wavelengths substantially simultaneously.

26. A scatterometer, comprising:

a lens configured to direct a radiation beam from a radiation source toward a pattern formed on a substrate; and

an edge adapted to be placed in one of opposite halves of an intermediate object plane associated with the lens,

wherein a property of the substrate is measured by a detector by measuring asymmetries between intensities of corresponding diffraction orders in an angle-resolved spectrum of a radiation beam diffracted from the pattern, at a plurality of angles and a plurality of wavelengths substantially simultaneously.

27. The scatterometer of claim 26, further comprising a non-polarizing beam splitter and a tilted mirror configured to couple off a portion of the radiation beam emitted from a radiation source for a separate measurement with the detector.

28. The scatterometer according to claim 26, wherein the edge is a Foucault knife edge.

29. The scatterometer according to claim 26, wherein a numerical aperture of the lens is at least 0.9.

30. The scatterometer according to claim 26, wherein a numerical aperture of the lens is at least 0.95.

31. The scatterometer according to claim 26, wherein wherein the detector is configured to measure at least one of (a) an intensity of a transverse magnetic and a transverse electric polarized light and (b) a phase difference between the transverse magnetic and the transverse electric polarized light.

32. A scatterometer, comprising:

a high numerical aperture lens configured to direct a radiation beam from a radiation source toward a substrate;

an edge adapted to be placed in one of opposite halves of an intermediate object plane associated with the high numerical aperture lens;

a wavelength multiplexer located between a radiation, source configured to provide the radiation beam, and the substrate; and

a wavelength demultiplexer located between the substrate and the detector located in a pupil plane of the high numerical aperture lens, and configured to measure the property of the reflected spectrum,

wherein a property of the substrate is measured by a detector by measuring an asymmetry in an angle-resolved spectrum of a radiation beam reflected from a surface of the substrate, at a plurality of angles and a plurality of wavelengths substantially simultaneously.