

LITHOGRAPHIC APPARATUS AND DEVICE MANUFACTURING METHOD

BACKGROUND

Field of the Invention

[0001] The present invention relates to a method to determine a feedforward transfer function of a control system in a lithographic apparatus, a lithographic apparatus and a method for manufacturing a device.

Description of the Related Art

[0002] A lithographic apparatus is a machine that applies a desired pattern onto a substrate, usually onto a target portion of the substrate. A lithographic apparatus can be used, for example, in the manufacture of integrated circuits (ICs). In such a case, a patterning device, which is alternatively referred to as a mask or a reticle, may be used to generate a circuit pattern to be formed on an individual layer of the IC. This pattern can be transferred onto a target portion (e.g. including part of, one, or several dies) on a substrate (e.g. a silicon wafer). Transfer of the pattern is typically via imaging onto a layer of radiation-sensitive material (resist) provided on the substrate. In general, a single substrate will contain a network of adjacent target portions that are successively patterned. Conventional lithographic apparatus include so-called steppers, in which each target portion is irradiated by exposing an entire pattern onto the target portion at once, and so-called scanners, in which each target portion is irradiated by scanning the pattern through a radiation beam in a given direction (the "scanning"-direction) while synchronously scanning the substrate parallel or anti-parallel to this direction. It is also possible to transfer the pattern from the patterning device to the substrate by imprinting the pattern onto the substrate.

[0003] As an example, in a so called step and scan lithographic apparatus, both reticle stage and wafer stage perform step- and scan moves at nanometer scale accuracy. Throughput demands require high velocity and acceleration levels. These high acceleration levels may be conflicting with a demand for position accuracy as dynamics of the stages are excited during acceleration phases. Vibrations need to settle in a certain settling time before exposure can

begin. To address these needs, sophisticated motion control is used to control 6 degrees of freedom (6 DOF) of mechanics of both reticle- and waferstage. Feedback control is used to guard stability and improve disturbance rejection, while feedforward control is used to achieve the desired tracking performance.

[0004] In the past years, iterative learning control has been considered to determine a feedforward signal in controlling a position related quantity (such as a position, a velocity, a jerk, etc) of a stage (such as a substrate stage or a reticle stage) of a lithographic apparatus. In a combined feedforward / feedback control system, iterative learning control provides for a learning of the feedforward signal by providing a setpoint stimulus to the control system, applying an initial or previously determined feedforward signal, measuring an error signal expressing a difference between a desired response of the stage and a measured or obtained response of the stage, and determining a following or subsequent feedforward signal from the error signal and the feedforward signal as applied. This process may be repeated iteratively, until a number of iterations has been achieved, or any other suitable criterion has been met, such as a convergence of the feedforward signal, a magnitude or norm of a remaining error signal, etc. Thus, in iterative learning control, an update is made to the feedforward signal off-line based on a measured error signal. After a few trials (iterations) a stable ILC converges to an ideal feedforward signal, which provides an ability to remove or at least reduce a repetitive tracking error .

[0005] Despite benefits that may be provided by application of iterative learning control, a limitation of it may be that such learned signal only holds for a particular setpoint (e.g. a setpoint versus time pattern corresponding to a stage movement pattern), being the setpoint that has been applied during the iterative learning. When the setpoint changes, system dynamics are excited differently. It is considered too time consuming, or even impossible to learn every possible occurring setpoint profile, because variations may be so numerous that almost infinitely many setpoint profiles may occur. This aspect is even aggravated in multi input, multi output systems, as there, a variety in setpoint profiles may increase as compared to a single input, single output system. Thus, it would be desirable to use the learned information from a limited amount of learned setpoints.

SUMMARY

[0006] It is desirable to enable an improvement in robustness of iterative learning control systems against setpoint variations.

[0007] According to an embodiment of the invention, there is provided a method to, in a lithographic apparatus, determine a feedforward transfer function of a control system, the method including a) iteratively learning by iterative learning control for a given setpoint signal a feedforward output signal of the control system; b) determining a relation between the learned feedforward output signal and the setpoint signal; and c) applying the relation as the feedforward transfer function of the control system.

[0008] In another embodiment of the invention, there is provided a lithographic apparatus including: a processing device to determine a feedforward transfer function of a control system, the processing device being provided with program instructions to: a) iteratively learn by iterative learning control for a given setpoint signal a feedforward output signal of the control system; b) determine a relation between the learned feedforward output signal and the setpoint signal; and c) apply the relation as the feedforward transfer function of the control system.

[0009] According to a further embodiment of the invention, there is provided a method for manufacturing a device, including: projecting by an irradiation beam a pattern from a patterning device onto a substrate; developing the irradiated substrate; and manufacturing a device from the developed substrate, wherein at least one of the patterning device and the substrate is held by a stage, a position of the stage being controlled by a control system having a feedforward transfer function, the feedforward transfer function being determined by: a) iteratively learning by iterative learning control for a given setpoint signal a feedforward output signal of the control system; b) determining a relation between the learned feedforward output signal and the setpoint signal; and c) applying the relation as the feedforward transfer function of the control system.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] 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:

[0011] Figure 1 depicts a lithographic apparatus according to an embodiment of the invention;

[0012] Figure 2 depicts a diagrammatic representation of a control system to iteratively learn a feedforward output signal;

[0013] Figure 3A and 3B depict a control diagram and a flow diagram respectively to explain iterative learning control;

[0014] Figure 4 depicts a diagrammatic representation of a control system including dual feedforward paths;

[0015] Figure 5A and 5B respectively depict a diagrammatic representation of a finite impulse response filter and a flow diagram illustrating actions to determine the finite impulse response or other relation;

[0016] Figure 6 depicts time diagrams of unfiltered errors and force of a stage positioning example;

[0017] Figure 7 depicts filter coefficients of a finite impulse response filter;

[0018] Figure 8 depicts a relation between remaining error and an order of the finite impulse response filter;

[0019] Figure 9 depicts a feedforwards of multi input, multi output control system; and

[0020] Figure 10 depicts a highly schematic control diagram of a feedforward forming part of an embodiment of the invention.

DETAILED DESCRIPTION

[0021] Figure 1 schematically depicts a lithographic apparatus according to one embodiment of the invention. The apparatus includes an illumination system (illuminator) IL configured to condition a radiation beam B (e.g. UV radiation or any other suitable radiation), a patterning device support structure (e.g. a mask table or mask support) MT constructed to support a patterning device (e.g. a mask) MA and connected to a first positioning device PM

configured to accurately position the patterning device in accordance with certain parameters. The apparatus also includes a substrate table (e.g. a wafer table) WT or "substrate support" constructed to hold a substrate (e.g. a resist-coated wafer) W and connected to a second positioning device PW configured to accurately position the substrate in accordance with certain parameters. The apparatus further includes a projection system (e.g. a refractive projection lens system) PS configured to project a pattern imparted to the radiation beam B by patterning device MA onto a target portion C (e.g. including one or more dies) of the substrate W.

[0022] The illumination system may include various types of optical components, such as refractive, reflective, magnetic, electromagnetic, electrostatic or other types of optical components, or any combination thereof, for directing, shaping, or controlling radiation.

[0023] The patterning device support structure (e.g., mask table or mask support) supports, i.e. bears the weight of, the patterning device. It holds the patterning device in a manner that depends on the orientation of the patterning device, the design of the lithographic apparatus, and other conditions, such as for example whether or not the patterning device is held in a vacuum environment. The patterning device support structure can use mechanical, vacuum, electrostatic or other clamping techniques to hold the patterning device. The patterning device support structure may be a frame or a table, for example, which may be fixed or movable as required. The patterning device support structure may ensure that the patterning device is at a desired position, for example with respect to the projection system. Any use of the terms "reticle" or "mask" herein may be considered synonymous with the more general term "patterning device."

[0024] The term "patterning device" used herein should be broadly interpreted as referring to any device that can be used to impart a radiation beam with a pattern in its cross-section so as to create a pattern in a target portion of the substrate. It should be noted that the pattern imparted to the radiation beam may not exactly correspond to the desired pattern in the target portion of the substrate, for example if the pattern includes phase-shifting features or so called assist features. Generally, the pattern imparted to the radiation beam will correspond to a particular functional layer in a device being created in the target portion, such as an integrated circuit.

[0025] The patterning device may be transmissive or reflective. Examples of patterning devices include masks, programmable mirror arrays, and programmable LCD panels. Masks are well known in lithography, and include mask types such as binary, alternating phase-shift, and attenuated phase-shift, as well as various hybrid mask types. An example of a programmable mirror array employs a matrix arrangement of small mirrors, each of which can be individually tilted so as to reflect an incoming radiation beam in different directions. The tilted mirrors impart a pattern in a radiation beam which is reflected by the mirror matrix.

[0026] The term "projection system" used herein should be broadly interpreted as encompassing any type of projection system, including refractive, reflective, catadioptric, magnetic, electromagnetic and electrostatic optical systems, or any combination thereof, as appropriate for the exposure radiation being used, or for other factors such as the use of an immersion liquid or the use of a vacuum. Any use of the term "projection lens" herein may be considered as synonymous with the more general term "projection system".

[0027] As here depicted, the apparatus is of a transmissive type (e.g. employing a transmissive mask). Alternatively, the apparatus may be of a reflective type (e.g. employing a programmable mirror array of a type as referred to above, or employing a reflective mask).

[0028] The lithographic apparatus may be of a type having two (dual stage) or more substrate tables or "substrate supports" (and/or two or more patterning device tables or "patterning device supports"). In such "multiple stage" machines the additional tables or supports may be used in parallel, or preparatory steps may be carried out on one or more tables or supports while one or more other tables or supports are being used for exposure.

[0029] The lithographic apparatus may also be of a type wherein at least a portion of the substrate may be covered by a liquid having a relatively high refractive index, e.g. water, so as to fill a space between the projection system and the substrate. An immersion liquid may also be applied to other spaces in the lithographic apparatus, for example, between the mask and the projection system. Immersion techniques can be used to increase the numerical aperture of projection systems. The term "immersion" as used herein does not mean that a structure, such as a substrate, must be submerged in liquid, but rather only means that a liquid is located between the projection system and the substrate during exposure.

[0030] Referring to Figure 1, the illuminator IL receives a radiation beam from a radiation source SO. The source and the lithographic apparatus may be separate entities, for example when the source is an excimer laser. In such cases, the source is not considered to form part of the lithographic apparatus and the radiation beam is passed from the source SO to the illuminator IL with the aid of a beam delivery system BD including, for example, suitable directing mirrors and/or a beam expander. In other cases the source may be an integral part of the lithographic apparatus, for example when the source is a mercury lamp. The source SO and the illuminator IL, together with the beam delivery system BD if required, may be referred to as a radiation system.

[0031] The illuminator IL may include an adjuster AD configured to adjust the angular intensity distribution of the radiation beam. Generally, at least the outer and/or inner radial extent (commonly referred to as σ -outer and σ -inner, respectively) of the intensity distribution in a pupil plane of the illuminator can be adjusted. In addition, the illuminator IL may include various other components, such as an integrator IN and a condenser CO. The illuminator may be used to condition the radiation beam, to have a desired uniformity and intensity distribution in its cross-section.

[0032] The radiation beam B is incident on the patterning device (e.g., mask MA), which is held on the patterning device support structure (e.g., mask table MT), and is patterned by the patterning device. Having traversed the patterning device MA, the radiation beam B passes through the projection system PS, which focuses the beam onto a target portion C of the substrate W. With the aid of the second positioning device PW and position sensor IF (e.g. an interferometric device, linear encoder or capacitive sensor), the substrate table WT can be moved accurately, e.g. so as to position different target portions C in the path of the radiation beam B. Similarly, the first positioning device PM and another position sensor (which is not explicitly depicted in Figure 1) can be used to accurately position the patterning device MA with respect to the path of the radiation beam B, e.g. after mechanical retrieval from a mask library, or during a scan. In general, movement of the patterning device table MT may be realized with the aid of a long-stroke module (coarse positioning) and a short-stroke module (fine positioning), which form part of the first positioning device PM. Similarly, movement of the substrate table WT or "substrate support" may be realized using a long-stroke module and a

short-stroke module, which form part of the second positioner PW. In the case of a stepper (as opposed to a scanner) the patterning device table MT may be connected to a short-stroke actuator only, or may be fixed. Patterning device MA and substrate W may be aligned using mask alignment marks M1, M2 and substrate alignment marks P1, P2. Although the substrate alignment marks as illustrated occupy dedicated target portions, they may be located in spaces between target portions (these are known as scribe-lane alignment marks). Similarly, in situations in which more than one die is provided on the patterning device MA, the mask alignment marks may be located between the dies.

[0033] The depicted apparatus could be used in at least one of the following modes:

[0034] 1. In step mode, the patterning device table MT or "patterning device support" and the substrate table WT or "substrate support" are kept essentially stationary, while an entire pattern imparted to the radiation beam is projected onto a target portion C at one time (i.e. a single static exposure). The substrate table WT or "substrate support" is then shifted in the X and/or Y direction so that a different target portion C can be exposed. In step mode, the maximum size of the exposure field limits the size of the target portion C imaged in a single static exposure.

[0035] 2. In scan mode, the patterning device table MT or "patterning device support" and the substrate table WT or "substrate support" are scanned synchronously while a pattern imparted to the radiation beam is projected onto a target portion C (i.e. a single dynamic exposure). The velocity and direction of the substrate table WT or "substrate support" relative to the patterning device table MT or "patterning device" may be determined by the (de-)magnification and image reversal characteristics of the projection system PS. In scan mode, the maximum size of the exposure field limits the width (in the non-scanning direction) of the target portion in a single dynamic exposure, whereas the length of the scanning motion determines the height (in the scanning direction) of the target portion.

[0036] 3. In another mode, the patterning device table MT or "patterning device support" is kept essentially stationary holding a programmable patterning device, and the substrate table WT or "substrate support" is moved or scanned while a pattern imparted to the radiation beam is projected onto a target portion C. In this mode, generally a pulsed radiation source is employed and the programmable patterning device is updated as required after each

movement of the substrate table WT or "substrate support" or in between successive radiation pulses during a scan. This mode of operation can be readily applied to maskless lithography that utilizes programmable patterning device, such as a programmable mirror array of a type as referred to above.

[0037] Combinations and/or variations on the above described modes of use or entirely different modes of use may also be employed.

[0038] A working principle of iterative learning control will now be described in a single input single output (SISO) control system framework. Iterative learning control may be abbreviated hereinafter as ILC. A control scheme including a feedback controller C controlling a process P such as a short-stroke patterning device stage, having a feedforward F and an ILC controller in a learning configuration, is depicted in Figure 2. An error signal e is determined as a difference between setpoint r and measured position y , as will be explained in more detail below. This error e is fed into feedback controller C. Feedforward controller F may in this example consists of a combination of acceleration and snap feedforward and provides a feedforward signal f_{ff} . The ILC controller computes a feedforward output signal, in this example a force signal f_{ilc} off-line based on previously measured errors. In operation, this force is added to a feedforward force from feedforward F, resulting in a more optimal feedforward operation. The learned force f_{ilc} has been learned with a given setpoint r , and therefore may provide a desired feedforward during operation when a setpoint r is applied which corresponds to the learned setpoint. In case however that another setpoint is applied, the learned feedforward may provide a suboptimal solution.

[0039] A principle of iterative learning control will now be explained in some more detail with reference to Figure 3A and 3B. Figure 3A depicts a so-called lifted ILC controller in an iterative (learning) domain. Herein e^k represents a matrix with error signal of a k th iteration, L represents a learning matrix, Z^{-1} a one iteration delay operator, I a unitary matrix and P_s matrix of the form:

$$\begin{pmatrix} h_0 & 0 & \Lambda & 0 \\ h_1 & h_0 & & \\ M & & O & \\ h_n & h_{n-1} & & h_0 \end{pmatrix}$$

Where parameters $h_0, \dots, h_n$ are so-called Markov parameters. Matrix P_s has a Toeplitz structure representing a process sensitivity dynamics of process P , and S a Toeplitz matrix of a sensitivity dynamics.

[0040] At each iteration the error is further reduced by learning a more appropriate ILC output signal to counteract errors. This can be expressed by the following error equations in a domain of iteration (also known as trial domain):

$$\begin{aligned} e^{k+1} &= S(r - d) - P_s f^{k+1} \\ e^{k+1} &= S(r - d) - P_s (f^k + L e^k) \\ e^{k+1} &= e^k - P_s L e^k \\ e^{k+1} &= (I - P_s L) e^k \end{aligned}$$

The above process is schematically depicted in Figure 3B showing a flow diagram where in 100 the setpoint signal is applied to the control system, thereby providing as the feedforward output signal an approximation of the feedforward output signal. Then, in 110 an error signal of the control system is recorded. Next, in 120 from the error signal and the approximation of the feedforward output signal a following or subsequent approximation of the feedforward output signal is determined. As indicated by a loop in Figure 3B, 100, 110 and 120 may be iteratively repeated applying each following or subsequent approximation of the feedforward output signal as the approximation of the feedforward output signal in a following or next repetition. After a predetermined criterion has been met (such as a number of iterations, a convergence criterium or any other suitable criterium), the approximation thereof which has been determined in a last repetition is provided as the feedforward output signal.

[0041] It is noted that the term setpoint or setpoint profile should be understood as a setpoint over a time period, e.g. a time sequence of setpoint signal values. In the example where

the control system includes a stage control system, the setpoint may include a profile representing a step and scan movement, e.g. including a sequence of acceleration, constant scanning speed, and deceleration. The setpoint may include a single quantity of a plurality of related quantities such as position, speed, acceleration etc.

[0042] The term feedforward output signal may also be understood as an output signal over time, e.g. a time sequence of feedforward signal values.

[0043] ILC may be used to converge to a feedforward output signal for a certain setpoint. This feedforward output signal is generally not optimal for other setpoints, as the thus learned feedforward output signal forms e.g. a time sequence of feedforward output values which have been learned to take account of errors occurring with the learned setpoint.

[0044] To address the above, application of a transfer function, such as a filter, as a transfer function between setpoint and ILC output signal is proposed. The transfer function may be determined from the learned ILC feedforward signal and the setpoint signal that has been applied during the learning of the ILC signal. Having determined such a relation, it allows not only for feedforward application given the setpoint that has been applied during the learning, however enables to adapt the feedforward signal towards other setpoint signals by determining the feedforward signal from the setpoint signal and the transfer function. Thereby, a sensitivity for setpoint variation may be reduced, whereas an implementation of ILC as compared to stored ILC output signal tables storing ILC output signal time sequences, may be achieved.

[0045] As an example, the relation may include a filter, such as a finite impulse response filter.

[0046] Other relations may be autoregressive (ARX), infinite impulse response (IIR), Wavelet or Fourier transformation type. An example of the above, applying as the relation a finite impulse response filter, will be provided below.

[0047] In generic terms, in a discrete time domain, the finite impulse response filter may consist of a plurality of coefficients: the coefficients, in cooperation with a plurality of setpoint values at different times in the time sequence of setpoint values, to provide the feedforward signal value at a given time.

[0048] In discrete time the above may be written as:

$$f_k^{ilc} = g_{-n1}r_{k+n1} + g_{-n1+1}r_{k+n1-1} + \Lambda + g_{n2}r_{k-n2}$$

where f_k represents the feedforward output signal at (sample) time k . r_{k+n1} represents the setpoint value at sample time $k+n1$, r_{k-n2} represents the setpoint value at sample time $k-n2$, and g_{-n1} represents filter coefficient $-n1$, while g_{n2} represents filter coefficient $n2$. It is noted that, in this example, mapping is non-causal, as f_k is a function of r_{k+n1} . In this example, a non-causal part of the filter is $n1$, whereas a causal part is $n2$. The above equation can also be written as a summation for a plurality :

$$f^{ilc}(k) = \sum_{i=-n1}^{n2} g(i)r(k-i)$$

In matrix form with R a Toeplitz matrix containing setpoint data, f a vector containing the learned feedforward signal, and g a vector with the filter coefficients:

$$\begin{pmatrix} f_{k1} \\ M \\ M \\ f_{k3} \end{pmatrix} = \begin{pmatrix} r_{k1+n1} & & r_{k1-n2} \\ M & O & \\ M & & r_{k1+n1} \\ r_{k2+n1} & & M \end{pmatrix} \begin{pmatrix} g_{-n1} \\ M \\ M \\ g_{n2} \end{pmatrix}$$

$f \qquad R \qquad g$

or,

$$f = R \cdot g$$

(*)

In simple terms, a set of equations now has been obtained providing as many more equations than unknown variables. This is because the setpoint values at the various sample times are known, while the iteratively learned feedforward signal is known too. Therefore, the unknown

variables, i.e. the coefficients of the filter, can be determined therefrom, with a reduction of noises contained in the data points.

[0049] Using matrix calculation, the above matrix equation may be solved using a linear least squares approximation based on normal equations:

$$g = (R^T R)^{-1} R^T f$$

Once coefficients $g(i)$ are determined, a generic relation between setpoint r and feedforward output signal has been determined. Preferably, $k \gg n$ (k is much larger than n), in other words a total number of discrete time setpoint output values of the setpoint is larger (more preferably significantly larger) than a number of coefficients in the determined relation. Therefore, apart from the possible benefits explained above, a data amount reduction may be achieved which may be beneficial from an implementation point of view.

[0050] The described process may be summarized, referring to Figure 5B, by: iteratively learning by iterative learning control for a given setpoint signal a feedforward output signal of the control system (block 200), determining a relation between the learned feedforward output signal and the setpoint signal (block 210); and applying the relation as the feedforward transfer function of the control system (block 220).

[0051] The relation that has been determined between the setpoint signal and the learned feedforward output signal, which in the above example forms a finite impulse response (abbreviated as FIR) filter, can be applied in a feedforward controller with a FIR filter F_{fir} parallel to a known feedforward F , as shown in Figure 4. By using dual feedforwards, accuracy may be improved as the FIR filter derived from the learned feedforward will function to reduce remaining errors from the known feedforward. FIR filters may be easy to program and they are inherently stable. In another example the FIR filter may be used as a sole feedforward.

[0052] The FIR filter F_{fir} may include a structure as schematically depicted in Figure 5A, with Z^{-1} represents a one sample time delay operator, Z represents an inverse thereof, i.e. a one sample time lead operator, and the coefficients g representing the coefficients of the finite impulse response filter.

[0053] When determining the relationship in a form of a FIR, an estimation of a number of coefficients can be made by deriving a time length (expressed in e.g. a number of sample times) from the iteratively determined feedforward output signal. When the iteratively learned output signal shows a response to a change in a value of the reference signal, the time length in number of sample times of the response can be derived therefrom and applied as a time length of the finite impulse response filter. In a discrete time domain, the time length may be expressed in a number of sample number of sample times (thus number of unit delay times of the unit delay operator depicted in Figure 5A) , providing a number of coefficients and thus a time length of the relation such as the finite impulse response filter. Also, it can be determined if the finite impulse response filter may be made acausal: in case that the iteratively learned feedforward output signal appears to anticipate a change in the setpoint signal, an acausal time length of this anticipation can be used to determine a number of coefficients expressing such acausality. As an example: presume that a step response in the setpoint signal results in an iteratively learned feedforward which provides a feedforward output signal starting to react to the step response 5 sample times before such step, and maintaining active 10 sample times after the step, then a relation may be defined having 15 coefficients forming a FIR filter having 5 acausal and 10 causal coefficients.

[0054] In case of a start-up of an acausal filter, setpoint values from the past would be needed, which may not be available. Zero's may be inserted for possibly unknown or unavailable setpoint signal values r from the past, filling the right upper triangular of R with zero's.

[0055] To be able to perform non-causal filtering, the setpoint r should be known a few samples ahead. This may be realized in practice by delaying the setpoint a few samples before feeding into the control loop, or by reading out a setpoint signal time sequence a few samples earlier and delaying it to provide the setpoint signal or reading out the sequence again some samples later. The FIR filter can use a time difference obtained thereby to become acausal. Note that, in case of a delay of the setpoint signal, such introduced delay is only encountered once during startup, and thus has no negative effects on e.g. a throughput of a lithographic apparatus equipped with such feedforward.

[0056] In an embodiment, iteratively learned feedforward output signals from different setpoints can be used to determine the FIR filter:

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix} (g)$$

The use of different setpoints as training data gives an average over different excitations making the filter more robust against setpoint variations.

[0057] Thus, in an embodiment use can be made of two (or more) feedforward output signals which are each iteratively learned for respective setpoint signals. Then, the relation is determined from the learned feedforward signals and the respective setpoint signals.

[0058] The process of learning a feedforward from several setpoints and use the learned data to train a FIR filter will be described below using an example. In this example, use is made of a Reticle Stage y- direction positioning control system. Three different setpoints with varying velocity and displacement values are used to train the filter:

	Start	End	Velocity	Acceleration	Jerk
Train Move 1	-0.12	0.12	2.2	69	8000
Train Move 2	-0.11	0.11	2.0	69	8000
Train Move 3	-0.08	0.08	1.6	69	8000
Test move 1	-0.1	0.1	1.8	69	8000

Filter performance is then evaluated with a test move at a setpoint that is different in both velocity and displacement from the training setpoints such that a proper validation is provided. These variations may cover possible variations that may occur in practice. Results are shown in Figure 6. The setpoint signal, in this example representing an acceleration of the stage, is represented by a discontinuous line. As a reference, an initial error is plotted with normal

feedforward only. An error achieved with an additional ILC feedforward signal after 15 iterations at the testing setpoint is indicated as well. Now, from the learned feedforward, an FIR filter is determined according to the above procedure. A tracking error after applying the FIR filter as a feedforward, is represented by "Error Fir" in Figure 15. It may be observed that the FIR filter realizes a performance, in other words, provides for a remaining error which is comparable to the remaining error that could have been achieved with the iteratively learned feedforward signal, be it that the FIR filter makes the feedforward more robust against variation in the setpoint signal, as the to be applied setpoint signal is fed to the FIR filter to calculate the feedforward output signal.

[0059] The trained filter coefficients g_i are plotted in Figure 7 with a non causal part $n_1 = 5$ and a causal part $n_2 = 10$, providing 15 coefficients in total.

[0060] When the FIR filter is trained at the same setpoint at which its performance is evaluated, an increasing filter order would result in better performance. A higher order (n th order) filter would then reproduce the data f_{ilc} of n samples exact. When however such high order filter is used at a setpoint which is different from the training setpoint, an optimum in filter order can be observed. Above this optimum the training data is overfitted which results in worse performance as data of the setpoint which is too far away in time to be relevant for the feedforward at a certain moment in time, is taken into account in the filter. This can be seen from Figure 8 where the RMS values of the measured error are plotted against filter order. Here, the FIR filter is trained at a fixed data set from three setpoints. The non-causal filter order is fixed to 5 samples. The causal filter order is increased from 5 to 400. An optimum where a lowest error is achieved, appears around 10 samples.

[0061] The above enables use of learned data in case the setpoint changes i.e. in the example of a lithographic apparatus stage positioning, when the stages perform different moves. Without being robust to setpoint variations, the use of ILC would be difficult, as for each possible setpoint signal, a feedforward would have to be learned. The concept described here proposes a way to handle setpoint variations with a generic FIR filter which can e.g. be added parallel to the conventional feedforward controller. Due to simple operations of the FIR filter, implementation may be relatively easy. Experiments have shown that substantially

zero-settling time of a tracking error can be realised in scanning direction of the stage. Thereby, throughput of the lithographic apparatus may be improved.

[0062] Further, to account for position dependent dynamics, a use of different position dependent FIR filters can be thought of. The filter coefficients themselves may become an explicit function of the position ($g(x,y)$). All filter coefficients are then updated during each sample time resulting in a position dependent filter. Interpolation between a limited amount of FIR filters is possible.

[0063] Although in the above, a single input, single output system has been described (e.g. a stage positioning in a single dimension), the concept described here may be applied in a multi input – multi output control system as illustrated with reference to Figure 9. In such multi input multi output control system, a plurality of (cross) feedforwards may have been provided to address cross effects such as cross talk, etc. From a setpoint generator signal, such as SPG-X, SPG-Y in the X respectively Y dimension, feedforward signals for an X, Y and Z dimension are obtained by respective feedforwards K_{ffxx} , K_{ffxy} , K_{ffxz} , and K_{ffyy} , K_{ffyx} , K_{ffyz} respectively. Such setup can be extended to as many variables as desired, whereby a number of cross feedforwards may increase with an increase in variables to be controlled (such as dimensions, rotations, positions, speeds, accelerations, etc.) In such systems, iterative learning control would provide for an adequate way to obtain feedforwards to compensate for such cross effects. However, numerous situations may occur because the multiple inputs and multiple outputs will result in many more different situations (combinations of setpoints, outputs, etc) to cope with. Therefore, many different situations would have to be learned which may make such system complex and difficult to implement. Because of the robustness against setpoint variations, the principle described in this document may provide adequate feedforward compensations here.

[0064] In a further embodiment, as described with reference to Figure 10, two feedforward output signals FF1, FF2 have been iteratively learned according to block 200 of Figure 5B for respective opposite polarities, directions (or other parameters of the setpoint signal, thereby, e.g. directional differences in a stage positioning system can be take into account). Processes shown in blocks 210 and 220 according to Figure 5B are performed for both feedforward output signals FF1, FF2 and their respective values of the setpoint signal R, to thereby provide respective feedforward transfer functions of the control system for

respective polarities of the setpoint signal. A selector may be provided to select either one of the feedforward signals FF1, FF2 depending on a polarity, direction or other parameter of the setpoint signal, thereby providing a different filter depending on the sign / polarity of the setpoint signal. Alternatively both filter outputs may be summed by a summation as indicated by Σ in Figure 10, thereby initialization effects that would occur in a case where an input of the filters would be switched (possibly resulting in initialization errors) may be largely prevented. Thus, a feedforward may be provided which is robust against setpoint variations, while taking account of the learned differences for the different setpoint signal polarities. Similarly such polarity may also be considered in a single acceleration phase, dependent on a polarity of a derivative of acceleration (jerk).

[0065] Different from the process of fitting a filter relation on the learned feedforward signal and the corresponding setpoint signal, alternatively the filter g can directly be determined from the measured error signal and the corresponding setpoint signal thereby omitting the learning process. The least squares problem is rewritten in terms of tracking error:

$$e = P_s \cdot R \cdot g$$

Herein matrix P_s is a Toeplitz matrix representing the process sensitivity dynamics.

[0066] As outlined above, the feedforward as described in this document may be implemented in a lithographic apparatus as for example described with reference to Figure 1, e.g. by having a processing device thereof provided with suitable program instructions to perform the described steps, thereby e.g. providing a control system to control a position quantity (e.g. a position, speed, acceleration, and/or jerk) of a stage such as the substrate stage WT or the patterning device table MT.

[0067] An embodiment of the invention may further be implemented in a device manufacturing method to manufacture a device such as a semiconductor device, processing unit, integrated circuit, electronic apparatus, etc. the method including projecting by an irradiation beam a pattern from a patterning device MA onto a substrate W; developing the irradiated substrate W; and manufacturing a device from the developed substrate,

wherein at least one of the patterning device MA and the substrate W is held by a stage (e.g. patterning device table MT and substrate table WT respectively), a position of the stage being controlled by a control system having a feedforward transfer function, the feedforward transfer function being determined by: a) iteratively learning by iterative learning control for a given setpoint signal a feedforward output signal of the control system (200); b) determining a relation between the learned feedforward output signal and the setpoint signal (210); and c) applying the relation as the feedforward transfer function of the control system (220).

[0068] Although embodiments of the invention have been described above in a context of lithography, and in particular in a context of stage positioning, any other accurate control application may benefit from the invention as described here.

[0069] Although specific reference may be made in this text to the use of lithographic apparatus in the manufacture of ICs, it should be understood that the lithographic apparatus described herein may have other applications, such as the manufacture of integrated optical systems, guidance and detection patterns for magnetic domain memories, flat-panel displays, liquid-crystal displays (LCDs), thin-film magnetic heads, etc. The skilled artisan will appreciate that, in the context of such alternative applications, any use of the terms "wafer" or "die" herein may be considered as synonymous with the more general terms "substrate" or "target portion", respectively. The substrate referred to herein may be processed, before or after exposure, in for example a track (a tool that typically applies a layer of resist to a substrate and develops the exposed resist), a metrology tool and/or an inspection tool. Where applicable, the disclosure herein may be applied to such and other substrate processing tools. Further, the substrate may be processed more than once, for example in order to create a multi-layer IC, so that the term substrate used herein may also refer to a substrate that already contains multiple processed layers.

[0070] Although specific reference may have been made above to the use of embodiments of the invention in the context of optical lithography, it will be appreciated that the invention may be used in other applications, for example imprint lithography, and where the context allows, is not limited to optical lithography. In imprint lithography a topography in a patterning device defines the pattern created on a substrate. The topography of the patterning

device may be pressed into a layer of resist supplied to the substrate whereupon the resist is cured by applying electromagnetic radiation, heat, pressure or a combination thereof. The patterning device is moved out of the resist leaving a pattern in it after the resist is cured.

[0071] The terms "radiation" and "beam" used herein encompass all types of electromagnetic radiation, including ultraviolet (UV) radiation (e.g. having a wavelength of or about 365, 248, 193, 157 or 126 nm) and extreme ultra-violet (EUV) radiation (e.g. having a wavelength in the range of 5-20 nm), as well as particle beams, such as ion beams or electron beams.

[0072] The term "lens", where the context allows, may refer to any one or combination of various types of optical components, including refractive, reflective, magnetic, electromagnetic and electrostatic optical components.

[0073] While specific embodiments of the invention have been described above, it will be appreciated that the invention may be practiced otherwise than as described. For example, the invention may take the form of a computer program containing one or more sequences of machine-readable instructions describing a method as disclosed above, or a data storage medium (e.g. semiconductor memory, magnetic or optical disk) having such a computer program stored therein.

[0074] The descriptions above are intended to be illustrative, not limiting. Thus, it will be apparent to one skilled in the art that modifications may be made to the invention as described without departing from the scope of the claims set out below.

WHAT IS CLAIMED IS:

1. A method for determining a feedforward transfer function of a control system in a lithographic apparatus, the method comprising:
 - a) iteratively learning a feedforward output signal of the control system using an iterative learning control for a given setpoint signal;
 - b) determining a relation between the learned feedforward output signal and the setpoint signal; and
 - c) applying the relation as the feedforward transfer function of the control system.
2. The method according to claim 1, wherein a) comprises
 - a1) applying the setpoint signal to the control system, thereby providing as the feedforward output signal an approximation of the feedforward output signal;
 - a2) recording an error signal of the control system;
 - a3) determining from the error signal and the approximation of the feedforward output signal a subsequent approximation of the feedforward output signal; and
 - a4) iteratively repeating a1), a2) and a3) applying each subsequent approximation of the feedforward output signal as the approximation of the feedforward output signal in a subsequent repetition; and
 - a5) providing as the feedforward output signal the approximation thereof which has been determined in a last repetition.
3. The method according to claim 1, wherein the relation comprises a finite impulse response filter.
4. The method according to claim 3, wherein a time length of the finite impulse response filter is derived from a time length of the learned feedforward output signal.

5. The method according to claim 1, wherein in a) at least two feedforward output signals are iteratively learned for two respective setpoint signals, and wherein in b) the relation is determined from the at least two iteratively learned feedforward signals and the respective setpoint signals.

6. The method according to claim 1, wherein the relation comprises an acausal relation.

7. The method according to claim 1, wherein two feedforward output signals are iteratively learned in a) for respective opposite polarities of the setpoint signal and wherein b) and c) are performed for both feedforward output signals and their respective setpoint signals so as to provide respective feedforward transfer functions of the control system for respective polarities of the setpoint signal.

8. The method according to claim 1, wherein the control system comprises a stage control system.

9. The method according to claim 1, wherein the control system comprises a multi input multi output control system, the feedforward transfer function of the control system including a cross feedforward transfer function in the multi input multi output control system.

10. A lithographic apparatus comprising:

a processing device to determine a feedforward transfer function of a control system, the processing device including program instructions to:

a) iteratively learn a feedforward output signal of the control system using an iterative learning control for a given setpoint signal;

b) determine a relation between the learned feedforward output signal and the setpoint signal; and

c) apply the relation as the feedforward transfer function of the control system.

11. The lithographic apparatus according to claim 10, wherein the processing device is configured such that a) is performed in accordance with a method comprising:

- a1) applying the setpoint signal to the control system, thereby providing as the feedforward output signal an approximation of the feedforward output signal;
- a2) recording an error signal of the control system;
- a3) determining from the error signal and the approximation of the feedforward output signal a subsequent approximation of the feedforward output signal; and
- a4) iteratively repeating a1), a2) and a3) applying each following approximation of the feedforward output signal as the approximation of the feedforward output signal in a subsequent repetition; and
- a5) providing as the feedforward output signal the approximation thereof which has been determined in a last repetition.

12. The lithographic apparatus according to claim 10, wherein the relation comprises a finite impulse response filter.

13. The lithographic apparatus according to claim 12, wherein the processing device is programmed to have a time length of the finite impulse response filter derived from a time length of the learned feedforward output signal.

14. The lithographic apparatus according to claim 10, wherein the processing device is programmed to iteratively learn in a) at least two feedforward output signals for two respective setpoint signals, and to determine in b) the relation from the at least two iteratively learned feedforward signals and the respective setpoint signals.

15. The lithographic apparatus according to claim 10, wherein the processing device is configured such that the relation comprises an acausal filter.

16. The lithographic apparatus according to claim 10, wherein the processing device is configured to have two feedforward output signals iteratively learned in a) for respective opposite polarities of the setpoint signal, and wherein b) and c) are performed for both feedforward output signals and their respective setpoint signals so as to provide respective feedforward transfer functions of the control system for respective polarities of the setpoint signal.

17. The lithographic apparatus according to claim 10, wherein the control system comprises a stage control system.

18. The lithographic apparatus according to claim 10, wherein the control system comprises a multi input multi output control system, the feedforward transfer function comprising a cross feedforward transfer function in the multi input multi output control system.

19. The lithographic apparatus according to claim 10, further comprising:
an illumination system configured to condition a radiation beam;
a patterning device support constructed to support a patterning device, the patterning device capable of imparting the radiation beam with a pattern in its cross-section to form a patterned radiation beam;
a substrate support constructed to hold a substrate; and
a projection system configured to project the patterned radiation beam onto a target portion of the substrate.

20. The lithographic apparatus according to claim 19, wherein the control system comprises a stage control system to control a position quantity of a stage, the stage comprising one of the supports.

21. A method for manufacturing a device, comprising:
patterning a beam of radiation with a patterning device;

projecting the patterned beam of radiation onto a substrate;
developing the substrate irradiated with the projected patterned beam of radiation; and
manufacturing a device from the developed substrate,
wherein at least one of the patterning device or the substrate is held by a stage, a position of the stage controlled by a control system including a feedforward transfer function, the feedforward transfer function determined by:

- a) iteratively learning a feedforward output signal of the control system using an iterative learning control for a given setpoint signal;
- b) determining a relation between the learned feedforward output signal and the setpoint signal; and
- c) applying the relation as the feedforward transfer function of the control system.

22. The method according to claim 21, wherein a) comprises

- a1) applying the setpoint signal to the control system, thereby providing as the feedforward output signal an approximation of the feedforward output signal;
- a2) recording an error signal of the control system;
- a3) determining from the error signal and the approximation of the feedforward output signal a subsequent approximation of the feedforward output signal; and
- a4) iteratively repeating a1), a2) and a3) applying each following approximation of the feedforward output signal as the approximation of the feedforward output signal in a subsequent repetition; and
- a5) providing as the feedforward output signal the approximation thereof which has been determined in a last repetition.

23. The method according to claim 22, wherein the relation comprises a finite impulse response filter.

24. The method according to claim 21, wherein a time length of the finite impulse response filter is derived from a time length of the learned feedforward output signal.

25. The method according to claim 21, wherein in a) at least two feedforward output signals are iteratively learned for two respective setpoint signals, and wherein in b) the relation is determined from the at least two iteratively learned feedforward signals and the respective setpoint signals.

26. The method according to claim 21, wherein the relation comprises an acausal relation.

27. The method according to claim 21, wherein two feedforward output signals have been iteratively learned in a) for respective opposite polarities of the setpoint signal, b) and c) being performed for both feedforward output signals and their respective setpoint signals so as to provide respective feedforward transfer functions of the control system for respective polarities of the setpoint signal.

28. The method according to claim 21, wherein the control system comprises a stage control system.

29. The method according to claim 21, wherein the control system comprises a multi input multi output control system, the feedforward transfer function comprises a cross feedforward transfer function in the multi input multi output control system.