Changhai Ru · Xinyu Liu
Yu Sun *Editors*

Nanopositioning Technologies

Fundamentals and Applications

Springer

Nanopositioning Technologies

Changhai Ru • Xinyu Liu • Yu Sun
Editors

Nanopositioning Technologies

Fundamentals and Applications

Springer

Editors
Changhai Ru
Harbin Engineering University
Harbin, China

Xinyu Liu
McGill University Mechanical Engineering
Montreal, Québec, Canada

Yu Sun
University of Toronto
Department of Electrical and Computing
Toronto, Ontario, Canada

ISBN 978-3-319-23852-4 ISBN 978-3-319-23853-1 (eBook)
DOI 10.1007/978-3-319-23853-1

Library of Congress Control Number: 2015960181

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Contents

Chapter 1
A Review of Stick–Slip Nanopositioning Actuators

Peng Pan, Feiyu Yang, Zhenhua Wang, Bowen Zhong, Lining Sun and Changhai Ru

Abstract There is a surge in the development of actuators based on the stick–slip of piezoelectric actuators in the past decade for the advantages of nanoscale resolution and long travel distances. These actuators, termed PSTA (piezoelectric stick–slip actuators), are widely used in chip assembly and cell manipulation. This chapter provides a comprehensive review of PSTA systems. Reported PSTAs are generalized with two specific principles which can be considered as a framework to further classify other PSTAs. In addition, the applications of PSTAs are categorized into two groups according to the number of degrees of freedom (DOF). This chapter also discusses driving signals to actuate different structures of PSTAs and control methods to achieve high positioning resolutions.

1.1 Introduction of Stick–Slip Nanopositioning

Nanopositioning actuators have been used in many applications. A number of applications need both nanoscale resolutions and large travel. For example, in the process of cell manipulation, it is critical to approach cells and make fine adjustments. This process requires large travel ranges and high resolution. In addition, actuators with large travel and high resolution are also effective to improve the efficiency of chip assembly. Thus, actuators having multi-scale motion ranges and high resolutions have tremendous potential. A growing demand for high precision and accuracy has spurred the development of new actuator technologies. Structures of actuators become simpler and smaller. This improvement enables several actuators to be combined into one system to achieve new functions.

A commonly observed phenomenon in dynamic systems is that motion is caused during the transition from static to sliding friction force. This phenomenon termed stick–slip can be employed to develop actuators. Actuators based on the stick–slip

P. Pan • F. Yang • Z. Wang • B. Zhong (✉) • L. Sun • C. Ru
Research Center of Robotics and Micro System & Collaborative Innovation Center of Suzhou Nano Science and Technology, Soochow University, Suzhou 215021, China
e-mail: panpeng1813@gmail.com; yangfeiyu001@163.com; wanghzenhua@suda.edu.cn; zhbw@suda.edu.cn; lnsun@hit.edu.cn; rchhai@gmail.com

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_1

principle have high accuracy, long travel range, and simpler mechanical structures [1]. Stick–slip actuators (STAs) often cooperate with piezoelectric actuators (PA). Piezoelectric actuators are composed of piezoelectric materials such as lead–PbZirconate Titanate (PZT), lead–PbMagnesium Niobate (PMN), and lead–PbZinc Niobate (PZN). Because of the inherent properties of piezoelectric materials, their positioning resolution is high but cannot be infinitely high. In addition, travel distances of piezoelectric actuators are rather small. Thus, these actuators typically provide a short travel distance with a high resolution. The combination of piezoelectric actuators and stick–slip actuators (i.e., piezoelectric stick–slip actuators) makes it possess a high resolution and a long travel distance [1].

Owning to these advantages, there is a surge in innovations of this type of actuators. Since the first piezoelectric stick–slip actuator was invented in 1976 [2], much research was done to advance the development of piezoelectric stick–slip actuators. Up to today, nearly 60 patents about piezoelectric stick–slip actuators have been disclosed, and a high number of papers have been published in the literature [3]. It is worth noting that conventional sawtooth signal is widely applied to actuate piezoelectric stick–slip actuators. However, this driving signal is not always effective to actuate piezoelectric stick–slip actuators of different structures. For example, conventional sawtooth signal is unsuitable for moving perpendicularly. In order to achieve better performance, many improvements have been made to construct new driving signals. In the meanwhile, control methods have also been applied to eliminate intrinsic errors. As progress is made, resolution and travel distance of piezoelectric stick–slip actuators can reach nanometer resolutions and centimeter travel distances, and speed can reach mm/s [4].

This review provides a comprehensive discussion of the principle, driving signals, control methods, and applications of piezoelectric stick–slip actuators. In Sect. 1.2, the principle of PSTA is discussed. The characteristics of PSTA structures are discussed in Sect. 1.3. Section 1.4 discusses a series of driving signals. In Sect. 1.5, control algorithms are compared and the performance of different PSATs is discussed. Applications based on the stick–slip principle are summarized in Sect. 1.6. In Sect. 1.7, conclusions and future directions are given.

1.2 PSTA Principle

A generalized view of stick–slip is provided in Fig. 1.1. Body A and Body B are initially at rest. F_e represents external force applied on Body B. F_f represents friction force exerted on Body B. When F_e is less than or equal to F_{fs} (maximum static friction force), Body A sticks to Body B, and displacement and speed of Body A are both zero relative to Body B. When F_e is greater than F_{fs}, Body B slips relative

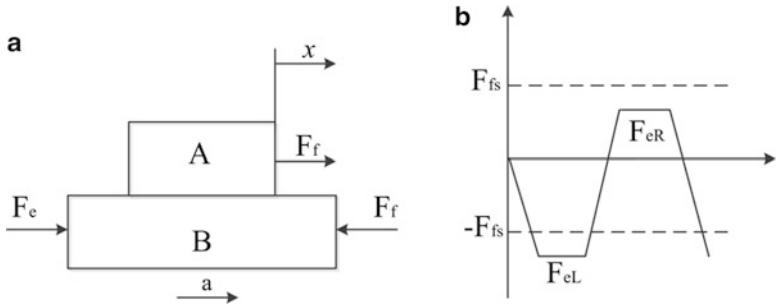


Fig. 1.1 Principle of stick–slip actuation. (**a**) Stick–slip phenomenon. (**b**) Change of external force

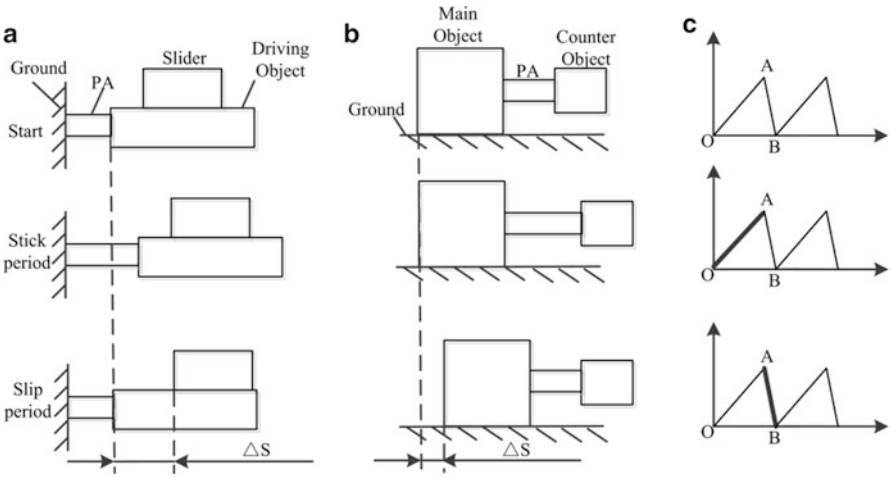


Fig. 1.2 Specific principles of PSTA. (**a**) Principle I. (**b**) Principle II. (**c**) Driving signal

to Body A. In dynamic systems, this phenomenon should be handled properly via controller design [5]. This phenomenon can also be utilized for actuation to achieve long travel distances and high resolutions. For example, Body A can move step by step along the right direction, relative to Body B by changing the magnitude and orientation of external force F_e. In particular, F_e in the right direction less than F_{fs} can make Body A sticks to Body B. F_e in the left direction greater than F_{fs} can make Body A slip to Body B. Therefore, by alternating the magnitude and orientation of F_e, Body A can move step by step along the right direction.

Based on this general principle, a number of PSTAs can be originated and can be classified into two groups according to the criterion of which type of force results in the actuating movement of the end effector. Figure 1.2 illustrates two specific PSTA principles that are discussed as follows.

1.2.1 Principle I

Figure 1.2a shows the first principle where the actuating movement of end effector is resulted from friction force in a cycle.

The three objects A, B, C, and D shown from left to right in this figure are known as ground (supporting frame), piezoelectric actuator (PA), driving object, and slider, respectively. PA is placed between ground and the driving object, and slider is regarded as end effector. The motion of PA is under control of applied voltages shown in Fig. 1.2c. The slider placed on the driving object can slide on the driving object with the aid of friction force. Actuation details of Principle I are summarized as follows.

Step 1: PA is in the initial state with no applied voltage.
Step 2: PA extends slowly with a small acceleration under the control of the first period of applied voltage. Driving object moves rightwards with the same acceleration. The inertial force applied to the slider depends on the mass of slider and the acceleration of PA and is less than the maximum static friction force. The slider sticks to the driving object and moves rightwards which is called "stick" period.
Step 3: PA retracts quickly with a large acceleration under the control of the second period of applied voltage. Driving object moves leftwards with the same acceleration. The inertial force applied to the slider depends on the mass of slider and the acceleration of PA and is larger than the maximum static friction force. The slider slips relative to the driving object and is at rest relative to the ground which is called "slip" period.
Step 4: PA retracts to the initial state when the voltage decreases to zero.

According to the four steps above, the slider achieves a step movement. Continuously repeating these four steps, the slider can move rightwards. If one wants to make the slider move leftwards, the shape of voltages should be reversed.

In this principle, the actuating movement of the slider occurs in the "stick" period. In this period friction force contributes to the actuating movement which means that friction force is directly related to actuating movement of end effector. In addition the stick period and slip period both occur between the slider and driving object. PSTA examples of this principle can be found in [6–8]. PSTA of this specific principle is further termed piezo friction driving actuator (PFDA).

1.2.2 Principle II

As shown in Fig. 1.2b, in this principle the actuating movement of the end effector results from inertial forces which depend on the mass of the counter object and the acceleration of PA in a cycle.

The three objects A, B, C, and D from left to right in this figure are known as the ground (supporting frame), main object, piezoelectric actuator (PA), and counter object, respectively. PA is placed between the main object and counter object, and the main object functions as the end effector. The motion of PA is under control of applied voltages shown in Fig. 1.2c. The main object placed on the ground can slide on the ground due to inertial forces. The working load can be applied to the main object. The counter object suspends in air and is not in contact with the ground. The actuation details of Principle II are as follows.

Step 1: PA is in the initial state with no applied voltage.
Step 2: PA extends slowly with a small acceleration under the control of the first period of applied voltage. Counter object moves rightwards with the same acceleration. The inertial force applied to the main object depends on the mass of counter object and the acceleration of PA and is less than the maximum static friction force from the ground. The main object sticks to the ground and is at rest relative to the ground which is called "stick" period.
Step 3: PA retracts quickly with a large acceleration under the control of the second period of applied voltage. The inertial force applied to the main object depends on the mass of counter object and the acceleration of PA and is larger than the maximum static friction force from the ground. The main object slips relative to the ground and moves rightwards relative to the ground which is called "slip" period.
Step 4: PA retracts to the initial state when the voltage decreases to zero.

According to these four steps, the main object achieves a step movement with respect to the ground. Continuously repeating these four steps, the main object can move rightwards together with the PA and the counter object. To make the slider move leftwards, the shape of applied voltages should be changed in a reversed order.

In this principle, the actuating movement of the main object occurs in the "slip" period. In this period the external force (i.e., inertial force) contributes to the actuating movement which means that friction force is not directly related to actuating movement of the end effector. In addition, the stick period and slip period both happen between the ground and main object. The PSTA examples of this principle can be found in [9–11]. The PSTA of this specific principle is further termed piezo inertial driving actuator (PIDA).

1.2.3 Discussion

In Principle I, the actuating movement of end effector is generated by the friction force. In Principle II, the actuating movement of end effector is generated by the inertial force. All PSTA can be classified into these two types. Some differences exist between these two principles. (1) Due to the limitation of maximum static friction force, the driving force of PSTA in the first principle is small. Differently, the driving force of PSTA in the second type of principle is large. (2) The resolution

of PFDA depends on PA and static friction force. The resolution of PIDA depends on PA and sliding friction force. However, the sliding friction force is more complex than static friction force. Without applying control algorithms, the accuracy of Principle I is higher than that of Principle II. (3) The travel distance of PSTA in the first principle is limited to the dimension of the driving object, whereas the travel distance of PSTA in the second principle has no limitation in theory.

1.3 Characteristics of PSTA Structure

Based on the two principles discussed above, many actuators have been developed in the past decades. These actuators can be grouped together according to the number of degrees of freedoms (DOF). In particular these actuators can be divided into two large groups, namely (1) those with one-DOF, and (2) those with multi-DOF.

1.3.1 Flexure Design

Flexure is typically used to transmit displacements without loss, resulting in a high displacement resolution of the motion. Sometimes the flexure is used to amplify motion to improve actuation stroke. An adjusting screw is often adopted to make a small deformation of flexure hinges. This small deformation of flexure hinge generates elastic potential energy which is released when the PA retracts. Thus, the released elastic potential energy produces a large inertial force which contributes to the decrease of backlash. Thus, the flexure is commonly used for transmitting motions produced by the piezoelectric stack actuator.

The typical flexures of one-DOF are the circular flexure hinge, shown in Fig. 1.3a and the right angle flexure hinge, shown in Fig. 1.3c. The rotation axis of the circular hinge is almost invariable, and the motion accuracy of it is high. Elliptical flexure hinge, derived from the circular hinge, is depicted in Fig. 1.3b. In order to obtain planar motion, some circular flexure hinges are combined [12]. However, the travel of planar motion is small which is not suitable for applications requiring large travels [13]. Right angle flexure hinge is a good alternative to the circular flexure hinge to achieve large travel. However, its motion accuracy is low for its deflection of the rotation axis [14]. Figure 1.3d shows a modified flexure hinge derived from the right angle flexure hinge termed corner-filleted flexure hinge [15]. The flexure hinges mentioned above only have one-DOF motion. To obtain multi-DOFs motions, spherical flexure hinge, having three axes, was developed and is shown in Fig. 1.4a. When the diameter of the thinner part of the spherical hinge increases to a certain degree, it only produces two-DOF motions, shown in Fig. 1.4b. It is worth noting that the stiffness of the thinner part is rather low, making it easier to break.

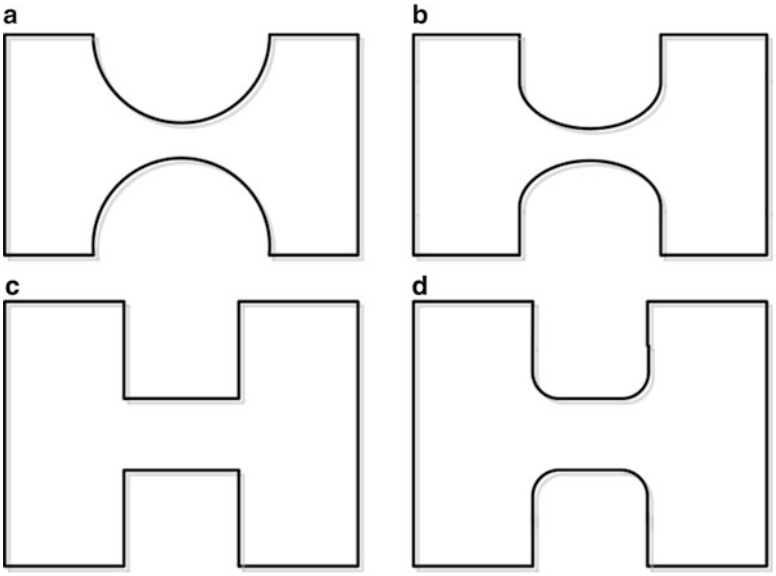


Fig. 1.3 Flexures of one-DOF. (**a**) Circular. (**b**) Elliptical. (**c**) Right angle. (**d**) Corner-filleted

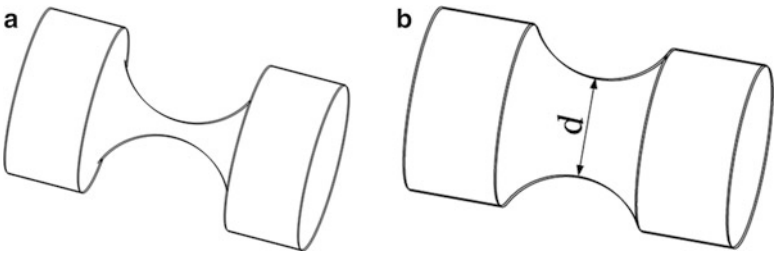


Fig. 1.4 Flexure of multiple DOFs. (**a**) Typical three-axis flexure: spherical flexure. (**b**) Prototype of a two axis flexure

1.3.2 Actuators with One-DOF

One-DOF actuators can be divided into two groups, one-DOF actuators with basic structure and one-DOF actuators with auxiliary structure. To improve the performance of this kind of actuators, auxiliary structure should be adopted.

1.3.2.1 One-DOF Actuators with Basic Structure

Among the actuators performing one-DOF actuation, some actuators only have four basic objects. The four basic objects referred to in the PFDA are PA, driving object, supporting frame, and slider (end effector), respectively. The four basic objects

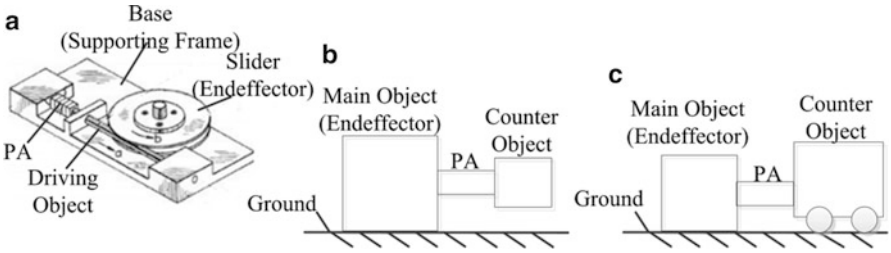


Fig. 1.5 Basic structures: (**a**) PFDA with basic structure [16], (**b**) PIDA with basic structure where counter object is suspended [9], (**c**) PIDA with basic structure where the counter object is supported [18]

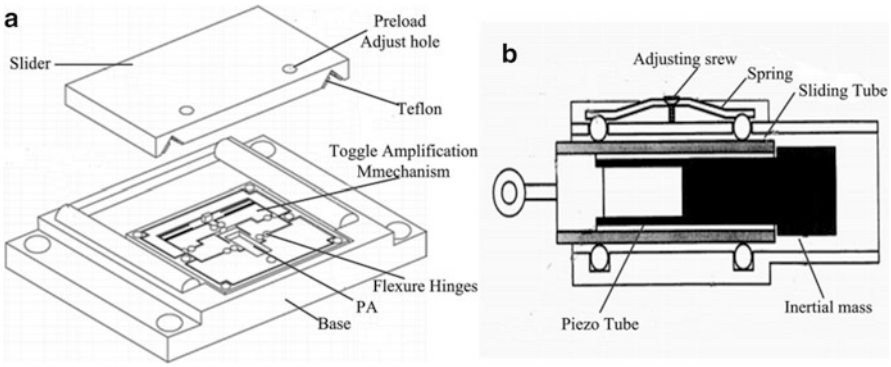


Fig. 1.6 (**a**) PFDA with auxiliary structures [21], (**b**) PIDA with auxiliary structures [37]

referred to in the PIDA are PA, main object (end effector), supporting frame, and counter object, respectively. The cases of PFDA have been referred to in [16, 17], shown in Fig. 1.5a. The cases of PIDA have been referred to in [9, 18], shown in Fig. 1.5b, c. The counter object shown in [9] is suspended, whereas the counter object shown in [18] is supported by the ground with a small friction force which is known as the supporting frame. Support by the ground with a small friction force makes PIDA more stable resulting in decrease of the lateral vibration of the counter object.

1.3.2.2 One-DOF Actuators with Auxiliary Structure

Due to manufacturing imperfections, fitting process, and the nature of the PSTA, it is difficult to achieve high performance by actuators with basic structures. It is necessary to adopt auxiliary structures to improve performance. Some auxiliary structure is applied to amplify the travel distance of PA to obtain large displacements and acceleration. Examples of amplification in PFDA are discussed in [19–22], and one of them is presented in Fig. 1.6a. For examples of amplification in PIDA, one can refer to [23]. Some auxiliary structures are adopted to adjust the friction

force or inertial force to reduce the backlash in the "slip" period and enable the vertical actuation of PSTA. For example, flexure hinge is applied to increase the inertial force [24–26]. Originally an adjusting screw is adopted to make a small deformation of the flexure hinge. This small deformation of flexure hinge generates elastic potential energy which is released when PA retracts. The released elastic potential energy produces a large inertial force which contributes to the decrease of backlash. Examples of adjusting mechanisms in PFDA can be found in [27–35]. Examples of adjusting mechanisms in PIDA can be found in [36–40].

1.3.3 Actuators with Multi-DOF

Multi-DOF actuators are divided into those with parallel mechanisms and those with serial mechanisms. Parallel mechanism means there is at least an end effector performing two or more DOFs. Serial mechanism means all the end effectors of multi-DOF actuators only perform one-DOF.

1.3.3.1 Multi-DOF Actuators with Parallel Mechanism

Among the actuators performing multi-DOF actuation, some end effectors of actuators can realize multi-DOF movement similar to the parallel robot. These kinds of multi-DOF actuators are called actuators with parallel mechanism, where a coupling phenomenon between actuation of some DOFs may exist as shown in Fig. 1.7. Multi-DOF PSTA with parallel mechanism can be found in [41] where PSTA can perform three DOF actuation including two linear actuations and a rotation actuation. Two linear actuations are called x-direction actuation and y-direction actuation, respectively. There is coupling between x-direction actuation and y-direction actuation. PFDA with parallel mechanism can be found in [42–46]. There is only one main object in PIDA which is recognized as the end effector. According to this characteristic, PIDA with multi-DOF are all actuators with parallel mechanism [9, 47, 48].

1.3.3.2 Multi-DOF Actuators with Serial Mechanism

If each end effector of actuators performs one-DOF actuation, the actuators are called multi-DOF actuators with serial mechanism which can be seen in Fig. 1.8. It means that there must be no coupling in this kind of actuators. Multi-DOF PFDA with serial mechanism was proposed in [49] where PSTA can perform three DOF actuations including two linear actuations and a rotation actuation. There are three end effectors, and each of them is actuated by its own set of PAs. Each end effector conducts one-DOF movement. PFDA with serial mechanism were also reported in [26, 50–52]. This concept is not suitable for PIDA. Therefore, there is no serial mechanism in PIDA.

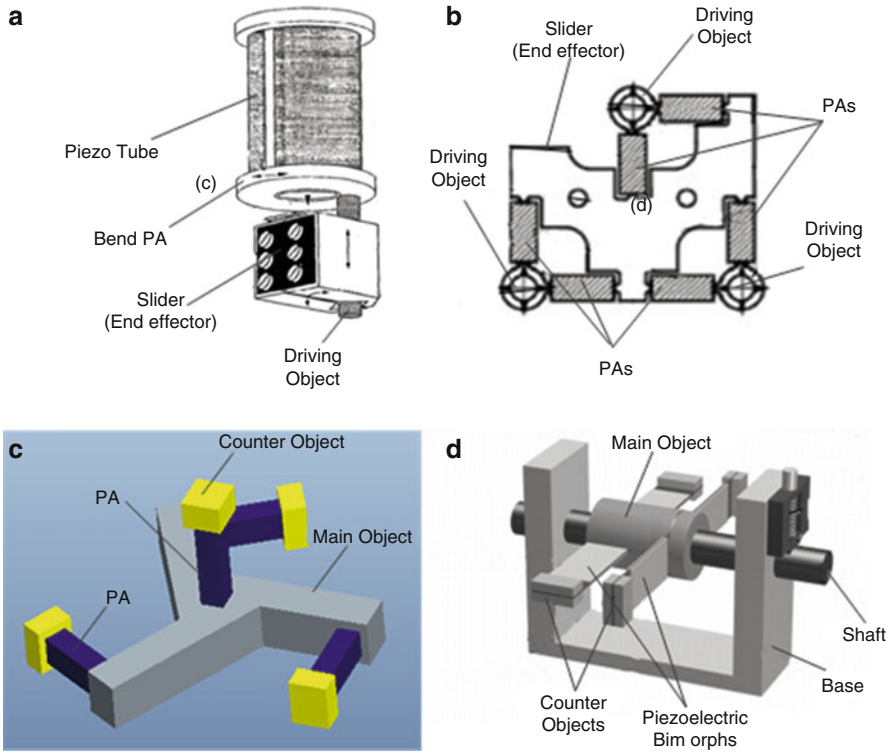


Fig. 1.7 Multi-PSTAs with parallel mechanism: (**a**) 2-DOF PFDAs with one vertical movement and a rotational movement [46], (**b**) 3-DOF PFDAs with two horizontal movements and a rotational movement [41], (**c**) 3-DOF PIDAs with two horizontal movements and a rotational movement [48], (**d**) 2-DOF PIDAs with one horizontal movement and a rotational movement [47]

1.4 Friction Model

The output displacement of PSTA is obtained by the change of friction force between the driving object and slider. Friction, affected by many factors, such as the relative sliding speed, acceleration, and displacement, is defined as the tangential forces between two contact surfaces. To obtain precise displacements, an appropriate friction model describing the characteristics of friction is critical. The friction model in the literature can be characterized into two categories: static friction model and dynamic friction model. The friction described in static friction model is the function of the relative speed. The friction is the function of the relative speed and displacement. Many frictional effects, including Column friction, viscous friction, static force, stribeck friction, pre-sliding displacement, varying static force, and frictional lag have been demonstrated which contribute to the accurate friction modeling.

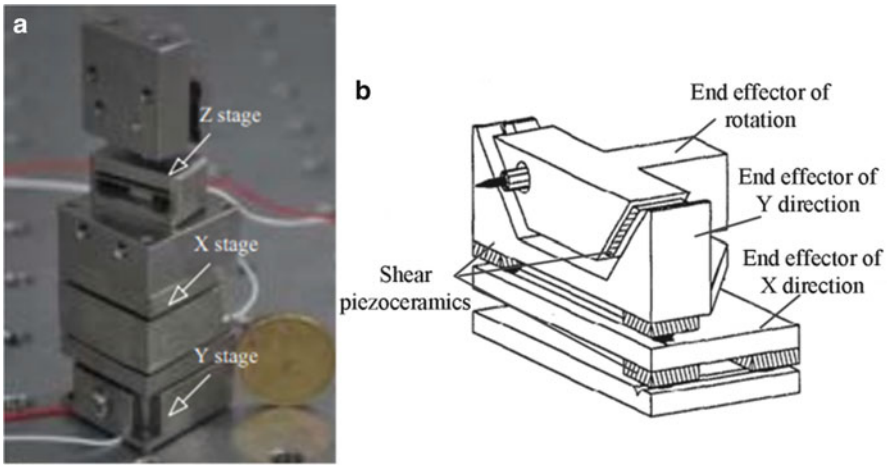


Fig. 1.8 Multi-PFDAs with serial mechanism: (**a**) 3-DOF PFDAs including 3 linear movements [26], (**b**) 3-DOF PFDAs including 2 linear movements and 1 rotational movement [49]

1.4.1 Static Friction Model

In the sixteenth century, Leonardo da Vinci proposed that the value of friction only depends on the normal load, and the direction of the friction is opposite to that of the direction of motion. Coulomb then presented the Column model based on da Vinci's conclusion. Figure 1.9a shows the characteristics of the Column model which indicates friction force is not associated with the contact area. In this figure, f_c represents the column friction. This model cannot describe friction when the relative speed is zero. A new model, called Column+viscous model, was developed via combining the Column model and the viscous model, as shown in Fig. 1.9b. This model indicates that friction is associated with the value of relative speed. However, this model also cannot describe the characteristics of friction when the relative speed is zero. In [53], static friction is considered to be associated with the external force when the relative speed is zero. It is experimentally demonstrated that the external force to actuate the object from zero to a speed is larger than zero and to keep the object sustained at this speed. Thus, the column friction is smaller than the maximum static friction. The model, representing the combination of Column, viscous, and static friction, is shown in Fig. 1.9c where f_s represents the maximum friction. Another typical model is demonstrated in Fig. 1.9d. In this model, friction force decreases with the increase of speed at relatively low speeds [15], which is defined as the stribeck effect. This model, integrating Column, viscous, static friction, and stribeck effect, gives an accurate description of the characteristics of friction at low speed and has been proven experimentally [54]. More detailed discussion about other friction models can be found in [55]. In the static friction model, it is widely believed that there is no relative movement between two contact

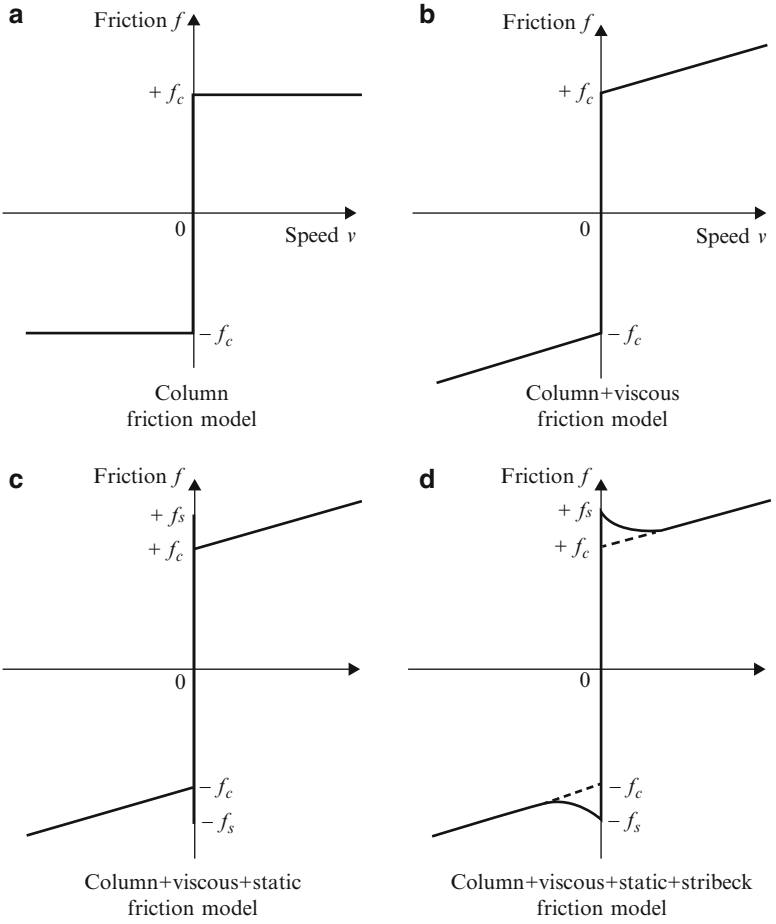


Fig. 1.9 Four typical static friction models [55]

surfaces when it is at the static friction stage. Although the static friction model can satisfy the requirements of many applications, it is not suitable for analyzing the principle of PSTA which requires the analysis of dynamic friction effects such as pre-sliding displacement, varying static force [56], and frictional lag [57]. Thus, more complicated dynamic friction models are introduced.

1.4.2 Dynamic Friction Model

From a mechanics standpoint, tangential contact and normal contact are both flexible by the external force. To accurately characterize the dynamic friction effect, a number of dynamic friction models have been established.

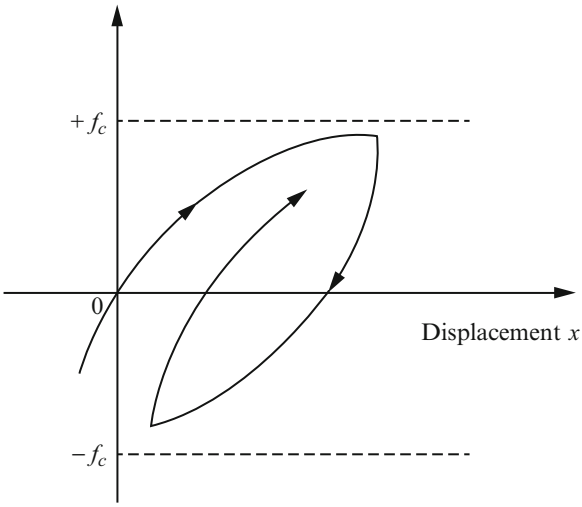


Fig. 1.10 Dahl friction model

1.4.2.1 Dahl Model

Dahl model is the simplest dynamic friction model which is the basis for other dynamic friction models. This model was proposed by Dahl [58, 59] after he conducted frictional experiments. Dahl discovered that contact peak between two contact surfaces results in a small flexible deformation, similar to that of a mechanical spring. Dahl formulated differential equations to characterize the relationships between displacement and friction before friction reached up to the maximum friction, shown in Fig. 1.10. In this model, the tangential compliance concept is utilized to introduce pre-sliding displacement, avoiding the discontinuity about the status switching in the static friction models. Although this model can characterize the pre-sliding displacement and the friction lag, the varying static force and the stribeck effect cannot be described by this model. Thus, this model is not suitable for the analysis of PSTA.

1.4.2.2 LuGre Model

In 1995, Canudas de Wit proposed the LuGre model which is based on the Dahl model [57]. This model is an extension of the Dahl model and is based on the bristle model [60]. The bristle model has an assumption that the contact surface is made up of a large amount of bristle which deforms randomly. Different from the bristle model which describes the random deformation of the bristle, the LuGre model is based on the average deformation of the bristle, shown in Fig. 1.11. The average deformation of the bristle is represented by the state variable z. The state-variable equation is

$$\frac{dz}{dt} = v - \frac{\sigma_0}{g(v)} z\,|v| \tag{1.1}$$

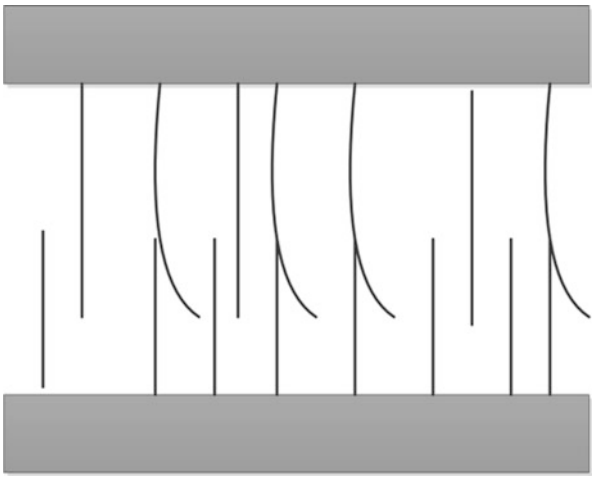


Fig. 1.11 Bristle model

where $g(v)$ describes the stribeck effect, and $g(v)$ is

$$g(v) = \frac{f_c + (f_s - f_c)\, e^{\left[-(|v|/v_s)^{\delta}\right]}}{\sigma_0} \tag{1.2}$$

Friction is generated by the deformation of the bristle, which can be described as a function of the average deformation of the bristle z and the relative speed v.

$$f = \sigma_0 z + \sigma_1 \frac{dz}{dt} + \sigma_2 v \tag{1.3}$$

where σ_0 is stiffness of the bristle, σ_1 is the micro damping coefficient, σ_2 is the viscous friction coefficient, f_s is the maximum static friction, f_c is the column friction, v_s is the stribeck speed, and δ is the empirical constant.

This model gives a comprehensive description of all friction effects including column friction, viscous friction, static force, stribeck friction, pre-sliding displacement, varying static force, and frictional lag. In addition, this model is continuous and can describe different friction states smoothly and continuously. There are six parameters in this model including two dynamic parameters and four static parameters. Compared with the static parameters, the dynamic parameters are difficult to be identified. In [61], the dissipation of this model was further analyzed.

1.4.2.3 Leuven Model

In [62], Swevers qualitatively compared simulation results based on the LuGre model and measured results of the shaft's pre-sliding motion of KUKA 361 IR. The results indicate that the LuGre model cannot accurately characterize friction

hysteresis. Thus, a more accurate model, called Leuven model, was developed by Swevers. This model also adopts the average deformation of the bristle.

$$\frac{dz}{dt} = v\left(1 - \text{sgn}\left(\frac{F_\text{d}(z)}{s(v) - F_\text{b}}\right)\left|\frac{F_\text{d}(z)}{s(v) - F_\text{b}}\right|^n\right) \tag{1.4}$$

and friction is given

$$f = F_\text{h}(z) + \sigma_1 \frac{dz}{dt} + \sigma_2 v \tag{1.5}$$

where n represents the coefficient describing the transition curve of different friction states, $s(v)$ represents the friction behavior at a constant speed, and $s(v)$ is

$$s(v) = \text{sgn}(v)\left(f_\text{c} + (f_\text{s} - f_\text{c})\, e^{-(|v|/v_\text{s})^\delta}\right) \tag{1.6}$$

$F_\text{h}(z)$ is the hysteresis friction force which is a static nonlinear force having non-local memory [15]. It consists of two parts.

$$F_\text{h}(z) = F_\text{b} + F_\text{d}(z) \tag{1.7}$$

where F_b represents the original position of the transition curve when the speed reverses, and $F_\text{d}(z)$ represents the transition curve generated in a certain time.

The Leuven model is not only capable of providing a more accurate description about the pre-sliding motion but also able to characterize the friction behavior that the LuGre model describes. There are six parameters and three mechanisms required to construct the Leuven model. In [63], a more detailed analysis and many modifications of the Leuven model are given.

1.4.3 Summary

The static model is not suitable to analyze PSTA because it cannot describe the pre-sliding motion, varying static friction, and friction lag. When the PA operates, generated displacements are small which is typically at micro- and nanoscales. This small displacement requires the concept of the bristle model for analysis. With the large variation of the PA's speed, the change of average deformation of the bristle starts to have influence on friction. LuGre model and Leuven model are appropriate for describing the friction behavior of PSTA. Although the Leuven model can give a more precise description of the friction behavior than the LuGre model, the difficulty of identifying six parameters and three mechanisms restricts its widespread use. The LuGre model is widely used for analyzing the friction behavior in the PSTA because it can provide a precise description of the friction behavior, and it is easier to be realized [4, 26, 64].

1.5 Driving Signals

In this section, different driving signals to actuate PSTA are introduced. As mentioned above, an inertial force larger than maximum static friction force should be obtained in order to achieve the "slip" period. To achieve the "stick" period, the inertial force should be smaller than maximum static friction force. In view of the change of inertial force during a cycle, the driving signal of PSTA should vary correspondingly. There are three types of driving signals utilized to actuate PSTA.

1.5.1 Sawtooth Waveform Driving Signal

Sawtooth waveform is widely used in PSTA. Many studies have been conducted on the sawtooth waveform to improve the performance of PSTA including step size, speed, positioning accuracy, and repeatability [27, 65–67]. These sawtooth waveforms can be classified into two types, including conventional sawtooth waveform (CSW) and modified sawtooth waveform (MSW).

1.5.1.1 Conventional Sawtooth Waveform

CSW has a slow ramp followed by a rapid retraction as shown in Fig. 1.12a. Anantheshwara et al. [67] studied a series of CSWs ranging from straight to gradually concave. Few studies have studied convex sawtooth waveforms [66]. The reason can be that a slower retraction where the turning point is not sharp and may make the performance of PSTA worse. In [65] different CSWs from concave CSW to straight CSW and all the way to convex CSW were studied. An unexpected result was shown that the convex CSW performs better compared to the concave CSW

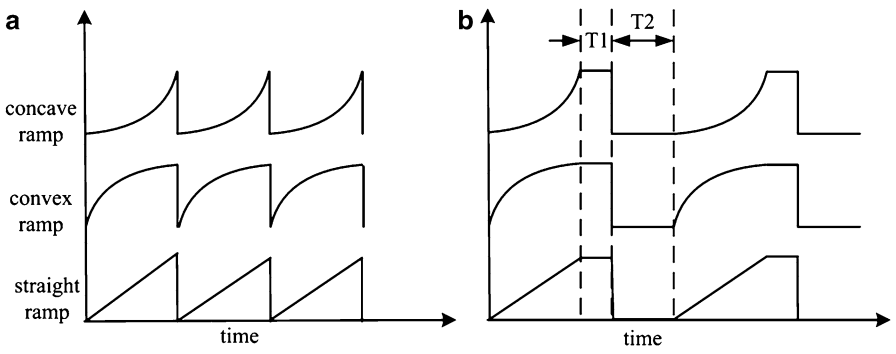


Fig. 1.12 Sawtooth waveform: (**a**) Conventional sawtooth waveform. (**b**) Modified sawtooth waveform

and straight CSW for the shear piezoceramics which performs as the source of mechanical energy. According to the simulation done in [65] with respect of the creeping nature of shear piezoceramics, a steeper retraction is gained in the convex CSW which can improve the performance of PSTA. Straight CSW and concave CSW are more suitable than convex CSW for the piezo tube slider because the creeping nature is comparatively small in the piezo tube slider.

1.5.1.2 Modified Sawtooth Waveform

Modifications have been made to CSW to enhance the performance of PSTA which leads to the development of MSW as shown in Fig. 1.12b. These modifications are described in [65] including a flat part of waveform right before retraction and a certain time delay right after retraction. One of the MSWs is shown in Fig. 1.12b where the straight ramp is followed by a flat segment of which time length is T1 before retraction and a time delay is T2 followed by the beginning of next period's straight ramp. These modifications are effective for shear piezoceramics as well as for piezo tube sliders. However, not all the values of T1 are effective for enhancing the performance of PSTA. The performance of PSTA becomes worse when T1 is too larger. With the optimum value of T1, the forward speed of sliding piece does not decrease significantly while the sliding piece of PA goes further forward with a distance d_1. When it retracts, the speed of sliding piece can generate a sufficient inertial force. Compared with the conventional straight SCW the step size of the MSW is larger. The forward speed of sliding piece decreases drastically which cannot generate enough inertial force to obtain the "slip" period when T1 is too large.

1.5.2 Square Waveform

Square waveform (SW) is a symmetrical driving signal as shown in Fig. 1.13. This driving signal is usually employed in PSTA of Principle II. There are four steps in a cycle of the square waveform driving signal.

Step 1: the square waveform signal is in the neutral state which generates no inertial force.
Step 2: the square waveform signal zooms to a definite value which generates a large inertial force.
Step 3: the square waveform signal remains unchanged which generates no inertial force.
Step 4: there is a deep descent to zero in the square waveform signal which generates a large inertial force in a reverse direction compared to that in step 2.

There are two large inertial forces which are reversed in a cycle of this signal. However, only one large inertial force is needed in Principle II. This signal is

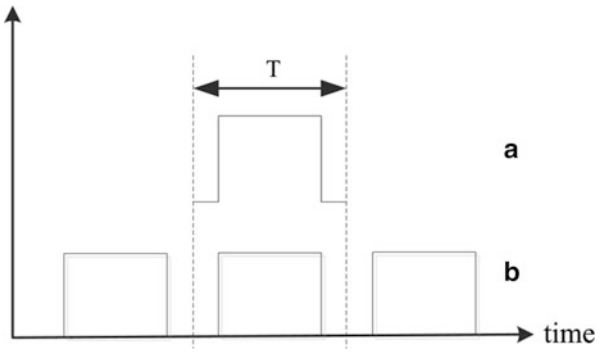


Fig. 1.13 Square waveform: (**a**) one cycle, (**b**) continuous step

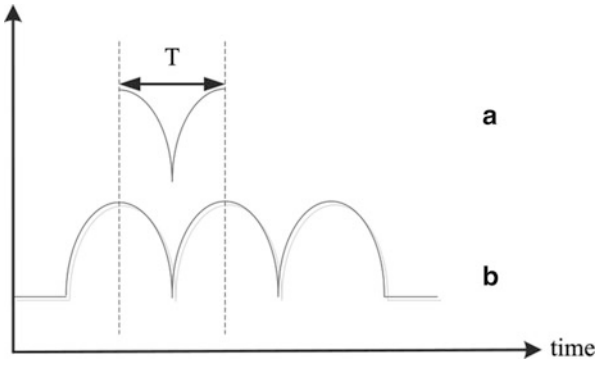


Fig. 1.14 Cycloid waveform: (**a**) one cycle, (**b**) continuous step

not suitable for PSTA of Principle II which contains one piezoelectric actuator. In fact, this signal is suitable for PSTA of Principle II including several piezoelectric actuators where some piezoelectric actuators are employed to adjust the friction force and some are employed to generate actuating movement [48].

Owning to the symmetry of the square waveform driving signal, the circuit is easy to design. In addition, this driving signal can result in a large driving force. However, this signal also can cause problems such as loud noises.

1.5.3 Cycloid Waveform (CW)

Cycloid waveform [68–70] driving signal is also a symmetrical signal as illustrated in Fig. 1.14. This signal is usually employed to actuate vertical PSTA. This is because the gravity of the end effector in vertical PSTA makes an impact on the movement. In particular when the end effector moves upwards, the acceleration of PA ($a_u < 0$) should satisfy $-u_s F_n - g < a_u$. When the end effector moves downwards, the acceleration of PA ($a_d > 0$) should satisfy $u_s F_n - g > a_d$. Thus, different

accelerations are required to obtain the desired movement of vertical PSTA's end effector. To obtain a desired step, only one of the inequalities should be satisfied during each cycle of signal. If both of them are satisfied during one cycle of signal, an erratic motion would appear. Hence, a sawtooth voltage is likely not suitable for the actuation of vertical PSTA against gravity because a sawtooth signal can lead to high accelerations of opposite amplitude at the turning points during each cycle. To realize the actuation of vertical PSTA, a cycloid signal is designed which is a signal with a quadratic time dependence. This cycloid signal can generate enough high acceleration in one direction while the acceleration remains small in the opposite direction. This performance can successfully actuate vertical PSTA upwards. To move the vertical PSTA downwards, the polarity of this cycloid signal can be inverted [31].

The cycloid signal is designed to actuate vertical PSTA which maintains the simplicity of the mechanical structure of PSTA. The performance of PSTAs gained by cycloid signal is high that nanoscale resolution can be obtained. However, the cycloid signal complicates circuit design. High-voltage broadband amplifiers are required to generate the cycloid signal [32].

1.6 Control Methods of PSTA

There are two modes of motion in PSTA [71] including stepping mode and scanning mode. The stepping mode means that PSTA moves step by step with a high velocity and high travel distance when the driving signal is applied to it repeatedly. The resolution of this mode is restricted to one step. When the difference between the actual position and target position is less than a step, the PSTA moves slowly until it reaches the target position. This is known as the scanning mode and the resolution of this mode is high. Therefore, two modes of control methods were developed for the PSTA. One kind of control method is implemented in the stepping method to achieve a high stroke positioning. Another kind is implemented in the scanning mode to achieve fine positioning [72].

1.6.1 Fine Positioning Control

Control is implemented in the scanning mode to achieve a fine position. These control methods are standard controllers such as PID, optimal controller, and robust controllers. These control methods are closed-loop.

1.6.2 High Stroke Positioning Control

This control is implemented in the stepping method. The open-loop control which is based on step counting is not suitable for this control mode. The values of every step are not identical with respect to the nonlinear phenomenon. In addition the external disturbances can lead to errors. Thus, closed-loop control methods were developed.

1.6.2.1 Basic Control Method

This control method is illustrated by the basic instruments in programming. When the resolution of sensor is better than a step, the basic instruments in programming are as follows:

$$\textit{While}\ |x_c - x| \geq \textit{step}\ \text{Do apply a step End While}$$

where x_c means the target position, x means the actual position.

When the resolution of the sensor is worse than a step, changes must be made to the algorithm in order to avoid oscillation. If the resolution is limited to the value of $n \times step$, it is

$$\textit{While}\ |x_c - x| \geq n \times \textit{step}\ \text{Do apply}\ n \times \textit{step}\ \text{End While}$$

This control method is easy to implement which can achieve the goal of reaching the target position.

1.6.2.2 Dehybridization Control Method (DCM)

This control method, based on the hybrid model of the stick–slip system with three discrete states and four continuous states, was proposed in [73]. The model of this control method is described in Fig. 1.15. The idea is described as follows:

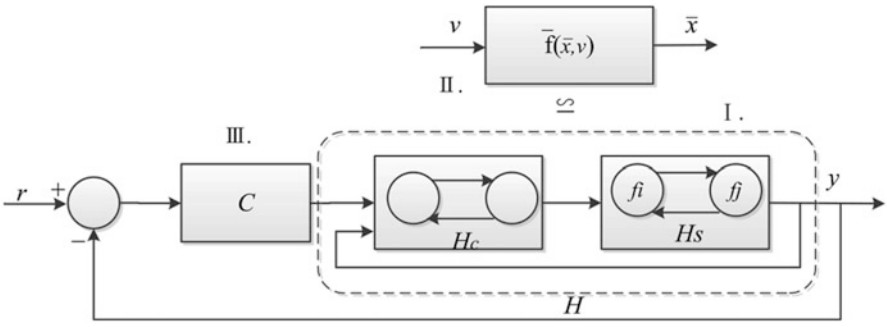


Fig. 1.15 Block diagram for the controller design via dehybridization [73]

(1) At first a continuous time model is adopted to approximate this hybrid system which is called dehybridization. (2) Then a standard continuous control method is applied to design the controller. A procedure named model reduction procedure was proposed which is related to the singular perturbation theory [74]. With the assistance of this procedure all states including discrete states and fast continuous states are eliminated. Because hybrid nature is removed in this procedure it is termed dehybridization. There are some prerequisites for this procedure. One of them is that there must be a repetitive mode in the hybrid system. However, the hybrid model of a stick–slip system does not possess a repetitive mode. An inner switching controller is applied to provide a repetitive mode to the hybrid model of stick–slip system. Then on the basis of feedback linearization a standard continuous control method is applied.

DCM includes three steps: (1) Providing a repetitive mode to hybrid system; (2) Dehybridization of the resulting repetitive hybrid system; and (3) A standard continuous control method is applied. By means of this control method there is no need to separate two control modes (stepping and scanning) to reach the target position. However, it is difficult to realize dehybridization. It is necessary to render a repetitive mode with the hybrid mode of stick–slip system which makes this control method more complex to implement.

1.6.2.3 Numerical Frequency Proportional Control Method (NFPCM)

The NFPCM was introduced in [75], the principle of which is described in Fig. 1.16. There are two modes of motion discussed above in PSTA. However, in this control method two modes of motion, stepping mode and scanning mode, are combined together in a simple way. The position error is converted into the clock signal of which the frequency is proportional to the position error. In accordance with the sign of position error, output signal of the counter increases or decreases. Then this signal is converted into the driving signal whose frequency is proportional to the position error and amplitude is constant. When PSTA reaches the target position, the frequency of driving signal is down to zero.

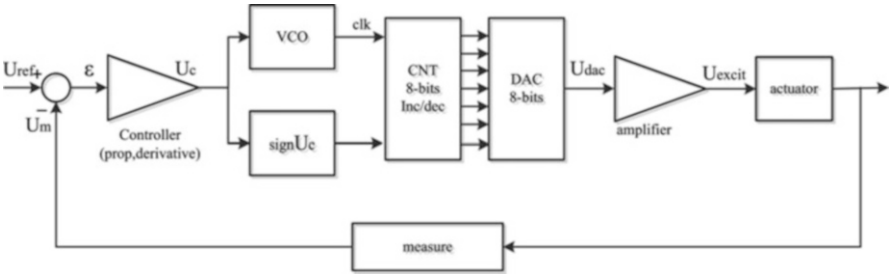


Fig. 1.16 Principle scheme of numerical frequency proportional control

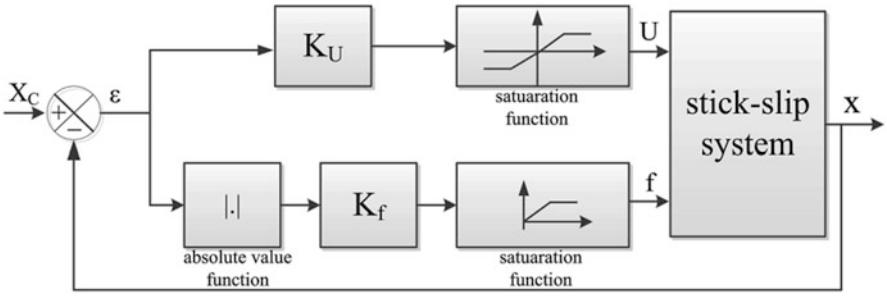


Fig. 1.17 Principle scheme of *U*/*f* proportional control

1.6.2.4 Voltage/Frequency Proportional Control Method (V/FPCM)

This control method was proposed in [72] encompassing three existing controllers which are sign controller, the frequency proportional controller, and the classical proportional controller, respectively. Figure 1.17 illustrates the principle of this control method. According to the proportional gains K_U, K_f and the absolute value of error $(X_c - X)$, there are five cases during displacement. When absolute value of the error $(X_c - X)$ is originally large, the values of voltage U and frequency f are both set to the state of saturation. In this phase the speed is constant. When the absolute value of the error $(X_c - X)$ becomes smaller, either the value of voltage U or the value of frequency f is not set to the state of saturation. The obtained speed in this phase is not constant which includes four cases.

There is a minimum voltage U_0 in this control method. When the amplitude U of the driving signal is below this minimum voltage U_0, there will be not enough torque and the "slip" period of PSTA will not be realized. Besides this parameter there is voltage saturation and frequency saturation in this control method to avoid overvoltages and restrict the PSTA to operate inside the linear zone of frequency. Two control modes (stepping and scanning) to reach the target position are not separated. With the increment of K_u, the static error is down to zero. When K_u increases, oscillation is generated. It is necessary to utilize a faster controller setup to make the implementation of a higher K_u possible. In this control method, the static error can be down to zero and the resolution can be higher than the senor resolution. The characteristics of different control methods are summarized in Table 1.1.

1.7 Applications in Industry

By virtue of its advantages including high resolution and large travel distance, PSTA is widely used in industry. This section discusses some of these examples.

Table 1.1 Characteristics of different control methods

Control method	Have been realized?	Two motion modes are separate?	Complexity	Resolution	Static error
BCM	Y	Y	Simple	Limited to a /n step	130 nm/10 μm
DCM	N	N	Complex	Unknown	Unknown
NFPCM	Y	N	A little complex	Limited to the resolution of sensor	Zero/50 nm
V/FPCM	Y	N	A little complex	Better than the resolution of sensor	Zero/10 μm

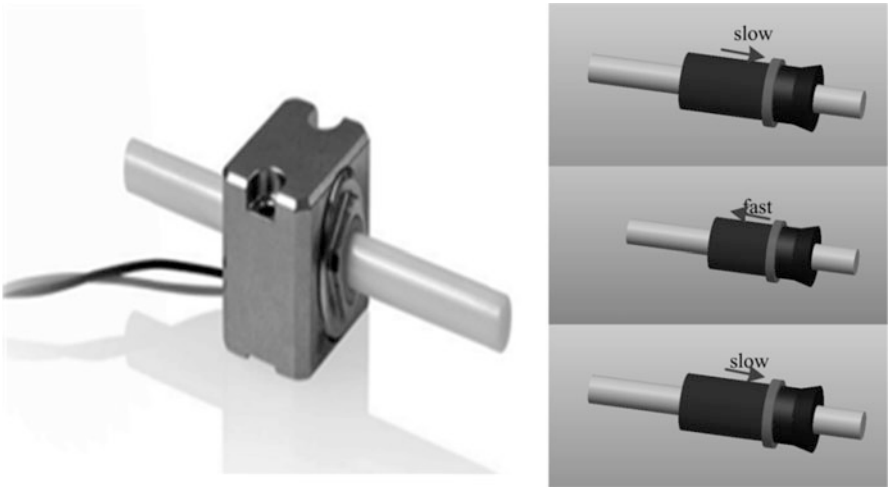


Fig. 1.18 PIShift drive designed by PI and the schematic of the PIShift

1.7.1 PhysikInstrumente (PI)

PI is a company that offers a broad portfolio of precision motion technologies including PSTA. One of the PSTAs designed by PI is shown in Fig. 1.18. This product is called PIShift drive which is on the basis of Principle II. Modified sawtooth driving signal is applied to one PA. The speed can reach more than 5 mm/s with a driving signal of above 20 kHz. This PSTA works silently at this high frequency and is self-locking when at rest. It is space-saving with relatively high holding forces of up to 10 N and has large travel distances.

1.7.2 SmarAct

SmarAct is a company that offers miniaturized positioners. The positioners offered by SmarAct are all based on the stick–slip principle, almost all of which are based on

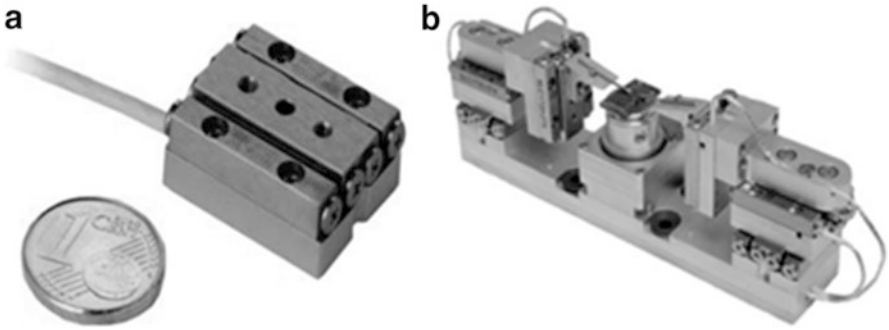


Fig. 1.19 (**a**) Linear PSTA designed by SmacAct, (**b**) a complete system designed by SmacAct

Principle I. The typical product of SmarAct is linear PSTA called SLC-1720 (http://www.smaract.de/) as shown in Fig. 1.19a. Their 7-DOF manipulator is a complete system consisting of 6 SLC-1720s and a rotary PSAT as shown in Fig. 1.19b. SLC-1720 is based on linear crossed-roller slides, and it is known for high rigidity and straightness. Therefore, it is ideally suitable for micro- or nanopositioning. The dimensions of SLC-1720 are $22 \times 17 \times 8.5$ mm^3 and the weight of it is about 13 g. It has a travel distance of 12 mm with a high resolution as much as 50 nm. The maximum speed of SLC-1720 can reach 13 mm/s with a high frequency of driving signal as much as 18.5 kHz. The 7-DOF manipulator is suitable for use inside the chamber of scanning electron microscope (SEM) and focused ion beam (FIB).

1.7.3 KlockeNanotechnik

KlockeNanotechnik offers a nanomotor (http://www.nanomotor.de/) shown in Fig. 1.20a to provide a linear motion driven by PA. It consists of a piezo tube, a cylindrical housing, and a slider in which a free axial hole exits. This free axial hole can be utilized to transport a tip, an electrode, or a microgripper. This nanomotor takes advantage of the stick–slip principle, having a positioning stroke up to centimeters with a maximum speed of up to 2 mm/s and a resolution down to the atomic scale. The driving force against gravity can reach more than 25 g. The length of it is half the size of a match stick shown in Fig. 1.20a. This nanomotor is suitable for ultra-vacuum, low temperature, and even underwater applications. The 3-DOF manipulator, shown in Fig. 1.20b, is one of the applications of the nanomotor. The manipulator comprises a titling table and several nanomotors, including a central nanomotor for z movement. A further advantage of this manipulator is that there is thermal compensation for the central nanomotor when it points downwards. The tip placed inside the titling table can be positioned in a range of $5 \times 5 \times 19$ mm^3.

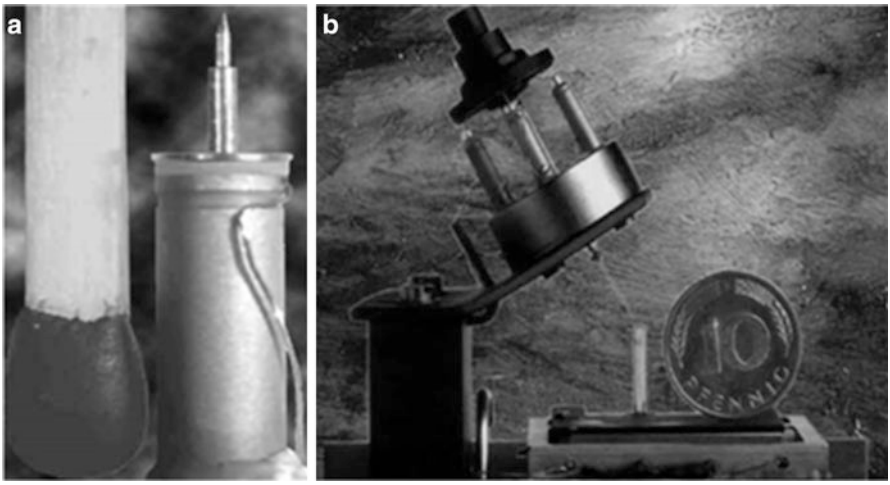


Fig. 1.20 (**a**) Nanomotor designed by KlockeNanotechnik, (**b**) a 3 DOF manipulator

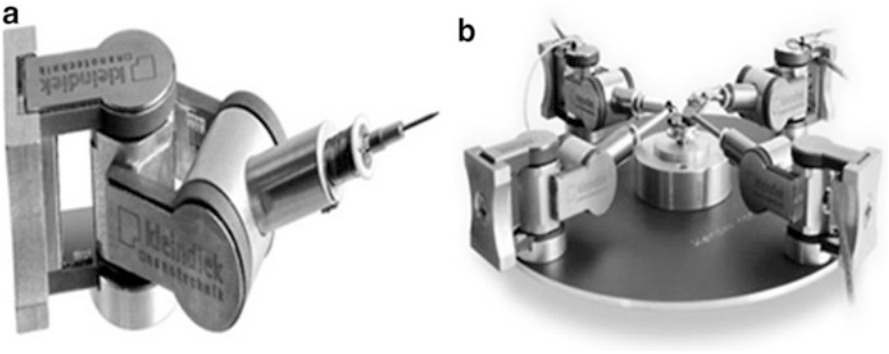


Fig. 1.21 (**a**) MM3A-EM designed by KleindiekNanotechnik, (**b**) a system consisted of several MM3A-Ems

1.7.4 KleindiekNanotechnik

Products designed by KleindiekNanotechnik (http://www.nanotechnik.com/mm3a-em.html) are based on the stick–slip principle. One of their manipulators is MM3A-EM. The dimensions of this manipulator are $60 \times 22 \times 25$ mm^3 which are easy to be integrated into electron microscopes. The manipulator consists of 3 PAs combined by serial mechanism. The first 2-dimensional motions are rotary motion with a speed of up to 10 mm/s and the resolutions of them are 5 nm and 3.5 nm, respectively. The last dimensional motion is linear motion with a speed of up to 2 mm/s, of which the resolution is 0.25 nm. The holding force of it can reach up to 1 N. The system shown in Fig. 1.21b consists of multiple of these manipulators.

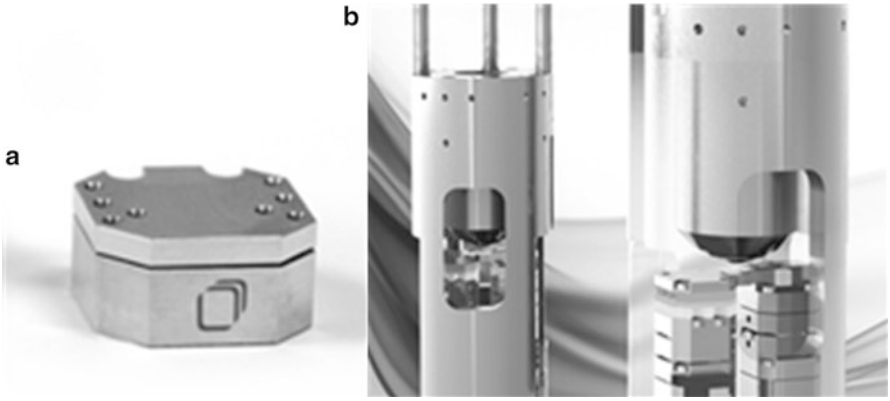


Fig. 1.22 (**a**) ANP_X101 designed by Attocube System. (**b**) AttoMICROSCOPY designed by Attocube System

1.7.5 Attocube System

The motion of PSTAs designed by Attocube System is reliable under environments such as high magnetic fields, ultra-high vacuum, and cryogenic temperatures. One of their typical PSTAs is ANP_X101 (http://www.attocube.com/) as shown in Fig. 1.22a. The dimensions of this PSTA are $24 \times 24 \times 11$ mm^3 with a weight of 20 g. When operating in the scanning mode, the scan range is 5 μm and the resolution can reach up to sub-nanometer. When operating in the stepping mode, the travel distance can be up to 5 mm with a minimum step size of 50 nm and a maximum speed of up to 3 mm/s. Maximum working load of this PSTA can be 1 N. AttoMICROSCOPY shown in Fig. 1.22b is a system that comprises multi-dimensional manipulators in the chamber of a microscope to perform physical manipulation tasks.

1.8 Conclusion

This chapter provided an overview of actuators based on stick–slip principle and piezoelectric actuators. The characteristics of different actuators based on stick-slip principle are summarized in Table 1.2. This kind of actuators is known for advantages of long travel distance and high resolution. Piezoelectric effect is the source of mechanical force, and the combination of inertial force and friction force is taken as the actuation force in PSTA. Existing PSTAs were categorized into two groups with two specific principles, namely PFDA and PIDA. The difference between the two groups is which type of force leads to the actual movement of the end effector. The applications of PSTAs are discussed on the basis of the number of DOFs, namely one-DOF actuators and multi-DOF actuators. Since conventional

Table 1.2 Characteristics of different PSTAs

Author	PR	DS	CM	RS	Size (mm)	WD (kg)	AM	P/S	DOF
Wen [48]	PIDA	SW	N/A	40	46 × 20 × 14	>1	Y	N/A	1-L(H)
Wen [48]	PIDA	SW	N/A	7.5 μrad	N/A	N/A	N	N/A	1-R
Wen [48]	PIDA	SW	BCM	30, 2.5 urad	N/A	>0.8	Y	Parallel	3-L(H), R
Pohl [27]	PFDA	CSW	N/A	40	30 × 30 × 20	1	N	N/A	1-L(H)
Niedermann et al. [76]	PFDA	CSW	N/A	10	N/A	0.12	N	N/A	1-L(H)
Renner et al. [31]	PFDA	CW	N/A	3	N/A	N/A	N	N/A	1-L(V)
Morita et al. [77]	PFDA	CSW	N/A	N/A	15(OD) × 10(ID) × 11	N/A	Y	N/A	1-R
Zhang et al. [78]	PFDA	CSW	N/A	26, 0.33 mrad	N/A	N/A	Y	Serial	2-L(H), R
Morita et al. [79]	PIDA	MSW	BCM	N/A	N/A	N/A	N	Parallel	3-L(H), R
Stieg et al. [66]	PFDA	CW	N/A	10	N/A	0.025	Y	N/A	1-L(V)
Agrait [29]	PFDA	CSW	N/A	0.5↑4↓	N/A	N/A	Y	N/A	1-L(V)
Breguet and Clavel [75]	PFDA	CSW	NFPCM	<5	N/A	N/A	N	N/A	1-L(H)
Rakotondrabe et al. [72]	PFDA	CSW	V/FPCM	<1.2	N/A	N/A	N	Parallel	2-L(H), R
Blackford et al. [46]	PFDA	CW	N/A	10	10 × 13 × 38	N/A	Y	Parallel	2-L(V), R
Howald et al. [45]	PFDA	CSW	N/A	4.8 urad	N/A	N/A	Y	Parallel	3-L(V), R
Libioulle et al. [42]	PFDA	CW	N/A	20,10↑30↓	30 × 30 × 25	N/A	Y	Parallel	2-L(H,V)
Tapson and Greene [51]	PFDA	CSW	N/A	20	35 × 35 × 35	0.05	N	Serial	2-L(H)
Göken [49]	PFDA	MSW	N/A	20	52 × 25 × 20	N/A	N	Serial	3-L(H), R
Yamagata and Higuchi [39]	PIDA	N/A	N/A	100	N/A	>1.3	N	N/A	1-L(H)
Darby and Pellegrino [37]	PIDA	MSW	N/A	<500	N/A	1	Y	N/A	1-L(H)
Yoshida et al. [18]	PFDA	MSW	N/A	N/A	150 × 35 × 15	0.15	Y	N/A	1-L(H)
Claeyssen et al. [23]	PIDA	N/A	N/A	<8	40 × 12 × 27	1	Y	N/A	1-L(H)

(continued)

Table 1.2 (continued)

Author	PR	DS	CM	RS	Size (mm)	WD (kg)	AM	P/S	DOF
PIShift drive (http://www.physikinstrumente.com/)	PIDA	MSW	N/A	300	N/A	1	N	N/A	1-L(H)
SLC-1720 (http://www.smaract.de/)	PFDA	CSW	N/A	50	$22 \times 17 \times 8.5$	4/0.1	Y	N/A	1-L(H/V)
Nanomotor (http://www.nanomotor.de/)	PFDA	N/A	N/A	0.1	$18 \times \Phi 33$	0.025	N	N/A	1-L(V)
MM3A-EM (http://www.nanotechnik.com/mm3a-em.html)	PFDA	N/A	N/A	5,3.5,0.25	$60 \times 22 \times 25$	0.1	N	Serial	3-L, R
ANP_X101 (http://www.attocube.com/)	PFDA	N/A	N/A	50	$24 \times 24 \times 11$	0.1	N	N/A	1-L(H)

PR principle, *DS* driving signal, *CM* control method, *RS* resolution of each actuator, *WD* working load, *AM* auxiliary mechanism, *P/S* parallel mechanism or serial mechanism

driving signals cannot meet the demand of actuation of different structures of PSTAs and achieving high performance, different driving signals were presented. The hysteresis of piezoelectric actuators and back-slash problem in PSTA affect their performance. Therefore, control methods were discussed. In order to achieve satisfactory performance, a number of issues must be considered For instance, the influence that driving signal poses on the performance of PSTA should be studied. Closed-loop control methods which are easy to implement should be explored to eliminate the intrinsic errors effectively. The mechanical structure of PSTA should be optimized including the improvement of auxiliary mechanisms to improve the dynamic oscillation of PSTA. Finally the thermal effect of PAs should be considered to reduce the adverse effect for the performance of PSTA.

References

1. P. Ouyang, R. Tjiptoprodjo, W. Zhang, G. Yang, Micro-motion devices technology: the state of arts review. Int. J. Adv. Manuf. Technol. **38**(5–6), 463–478 (2008)
2. A.L. Soderqvist, U.S. Patent 3,957,162, 1976
3. Z. Zhang, Q. An, J. Li, W. Zhang, Piezoelectric friction–inertia actuator—a critical review and future perspective. Int. J. Adv. Manuf. Technol. **62**(5–8), 669–685 (2012)
4. B.W. Zhong, Research on key technologies of cross-scale precision platform based on stick–slip driving. Dissertation, Harbin Institute of Technology, 2012
5. W.S. Owen, E.A. Croft, The reduction of stick–slip friction in hydraulic actuators. IEEE/ASME Trans. Mechatron. **8**(3), 362–371 (2003)
6. J. Peng, X. Chen, Modeling of piezoelectric-driven stick–slip actuators. IEEE/ASME Trans. Mechatron. **16**(2), 394–399 (2011)
7. J. Li, G. Yang, W. Zhang, S. Tu, X. Chen, Thermal effect on piezoelectric stick–slip actuator systems. Rev. Sci. Instrum. **79**(4), 046108 (2008)
8. Q. Zhang, X. Chen, Q. Yang, W. Zhang, Development and characterization of a novel piezoelectric-driven stick–slip actuator with anisotropic-friction surfaces. Int. J. Adv. Manuf. Technol. **61**(9–12), 1029–1034 (2012)
9. T. Higuchi, M. Watanabe, U.S. Patent 4,894,579, 1990
10. J.L. Ha, R.F. Fung, C.S. Yang, Hysteresis identification and dynamic responses of the impact drive mechanism. J. Sound Vib. **283**(3), 943–956 (2005)
11. Y.-T. Liu, T. Higuchi, R.-F. Fung, A novel precision positioning table utilizing impact force of spring-mounted piezoelectric actuator—part II: theoretical analysis. Precis. Eng. **27**(1), 22–31 (2003)
12. Z.J. Du, W. Dong, L.N. Sun, The flexure hinge and its application in the precision parallel robot. J. Harbin Inst. Technol. **38**(9), 1469–1473 (2006)
13. J.M. Paros, How to design flexure hinges. Mach. Des. **37**, 151–156 (1965)
14. F.S. Eastman, The design of flexure pivots. J. Aeronaut. Sci. (Inst. Aeronaut. Sci.) **4**(1) (2012)
15. B. Armstrong-Helouvry, *Control of Machines with Friction*, vol 128 (Springer Science & Business Media, 1991)
16. K.-J. Lim, J.-S. Lee, S.-H. Park, S.-H. Kang, H.-H. Kim, Fabrication and characteristics of impact type ultrasonic motor. J. Eur. Ceram. Soc. **27**(13), 4159–4162 (2007)
17. Y. Okamoto, Y. Tanijiri, U.S. Patent 6,092,431, 2000
18. R. Yoshida, Y. Okamoto, T. Higuchi, A. Hamamatsu, Development of smooth impact drive machine—proposal of driving mechanism and basic performance (in Japanese). J. Jpn. Soc. Precis. Eng. **65**(1), 111–115 (1999)

19. S. Chang, S. Li, A high resolution long travel friction-drive micropositioner with programmable step size. Rev. Sci. Instrum. **70**(6), 2776–2782 (1999)
20. Y. Wang, S. Chang, Design and performance of a piezoelectric actuated precise rotary positioner. Rev. Sci. Instrum. **77**(10), 105101 (2006)
21. C.-L. Chu, S.-H. Fan, A novel long-travel piezoelectric-driven linear nanopositioning stage. Precis. Eng. **30**(1), 85–95 (2006)
22. C. Belly, MRI-compliant piezo micro-actuator (2009), http://www.lirmm.fr/uee09/doc/Student_s_%20presentations/Christian_Belly.pdf. Accessed 7 Jan 2010
23. F. Claeyssen, A. Ducamp, F. Barillot, R. Le Letty, T. Porchez, O. Sosnicki, C. Belly, Stepping piezoelectric actuators based on APAs, in *Proceedings of the Actuator*, 2008
24. B.B. Nie, Research on precision rotary positioner based on stick–slip driving. Dissertation, Harbin Institute of Technology, 2013
25. Z.S.R.W.T. Guoan, Y.M.S. Lining, Designing and dynamic modeling of 1D nanopositioner based on stick–slip motion principle. J. Mech. Eng. **19**, 004 (2012)
26. W. Rong, S. Zhang, M. Yu, L. Sun, A 3D stick–slip nanopositioner for nanomanipulation, in *International Conference on Mechatronics and Automation (ICMA)* (IEEE, 2011), pp. 195–199
27. D.W. Pohl, Dynamic piezoelectric translation devices. Rev. Sci. Instrum. **58**(1), 54–57 (1987)
28. Y. Okamoto, R. Yoshida, Development of linear actuators using piezoelectric elements. Electron. Commun. Jpn. (Part III: Fundamental Electronic Science) **81**(11), 11–17 (1998)
29. N. Agrait, Vertical inertial piezoelectric translation device for a scanning tunneling microscope. Rev. Sci. Instrum. **63**(1), 263–264 (1992)
30. I. Altfeder, A. Volodin, Low-temperature scanning tunneling microscope with a reliable piezoelectrical coarse approach mechanism. Rev. Sci. Instrum. **64**(11), 3157–3160 (1993)
31. C. Renner, P. Niedermann, A. Kent, A vertical piezoelectric inertial slider. Rev. Sci. Instrum. **61**(3), 965–967 (1990)
32. V. Yakimov, Vertical ramp-actuated inertial micropositioner with a rolling-balls guide. Rev. Sci. Instrum. **68**(1), 136–139 (1997)
33. R. Yoshida, Y. Okamoto, K. Ishibashi, Y. Tanijiri, H. Okada, Movable stage utilizing electromechanical transducer. Google Patents, 1998
34. Q.-h. Lu, Z.-j. Gao, G.-z. Yan, D.-t. Yan, The study on miniature inertial piezo-actuators. Piezoelectrics Acoustooptics **2**, 009 (2004)
35. Sasaki, U.S. Patent 0,036,533 A1, 2007
36. T. Jiang, T. Ng, K. Lam, Optimization of a piezoelectric ceramic actuator. Sens. Actuators A: Phys. **84**(1), 81–94 (2000)
37. A. Darby, S. Pellegrino, Inertial stick–slip actuator for active control of shape and vibration. J. Intell. Mater. Syst. Struct. **8**(12), 1001–1011 (1997)
38. H.-j. Zhang, F. Huang, Piezo impact drive mechanism for precise approach and manipulation. J. Zhejiang Univ. Eng. Sci. **34**(5), 519–522 (2000)
39. Y. Yamagata, T. Higuchi, A micropositioning device for precision automatic assembly using impact force of piezoelectric elements, in *1995 IEEE International Conference on Robotics and Automation (ICRA)* (1995) pp. 666–671
40. T. Higuchi, K. Furutani, Y. Yamagata, K.-i. Kudoh, M. Ogawa, Improvement of velocity of impact drive mechanism by controlling friction. J. Jpn. Soc. Precis. Eng. Seimitsu Kogaku Kaishi **58**(8), 1327–1332 (1992)
41. J.M. Breguet, Actuionneurs "stick and slip" pour micromanipulateurs. Dissertation, École Polytechnique Fédérale de Lausanne, 1998
42. L. Libioulle, A. Ronda, I. Derycke, J. Vigneron, J. Gilles, Vertical two-dimensional piezoelectric inertial slider for scanning tunneling microscope. Rev. Sci. Instrum. **64**(6), 1489–1494 (1993)
43. A. Bergander, W. Driesen, A. Lal, T. Varidel, M. Meizoso, H. Bleuler, J.-M. Breguet, Position feedback for microrobots based on scanning probe microscopy, in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2004), pp. 1734–1739

44. B. Blackford, M. Jericho, Simple two-dimensional piezoelectric micropositioner for a scanning tunneling microscope. Rev. Sci. Instrum. **61**(1), 182–184 (1990)
45. L. Howald, H. Rudin, H.J. Güntherodt, Piezoelectric inertial stepping motor with spherical rotor. Rev. Sci. Instrum. **63**(8), 3909–3912 (1992)
46. B. Blackford, M. Jericho, M. Boudreau, A vertical/horizontal two-dimensional piezoelectric driven inertial slider micropositioner for cryogenic applications. Rev. Sci. Instrum. **63**(4), 2206–2209 (1992)
47. H. Zhang, G. Cheng, H. Zhao, P. Zeng, Z. Yang, Two dimensional precise actuator driven by piezoelectric bimorph impact. J. Jilin Univ. (Eng. Technol. Ed.) **36**(1), 67–71 (2006)
48. J.M. Wen, Study on planar inertia piezoelectric moving mechanism. Dissertation, Jilin University, 2009
49. M. Göken, Scanning tunneling microscopy in UHV with an X, Y, Z micropositioner. Rev. Sci. Instrum. **65**(7), 2252–2254 (1994)
50. R. Brockenbrough, J. Lyding, Inertial tip translator for a scanning tunneling microscope. Rev. Sci. Instrum. **64**(8), 2225–2228 (1993)
51. J. Tapson, J. Greene, A simple dynamic piezoelectric X-Y translation stage suitable for scanning probe microscopes. Rev. Sci. Instrum. **64**(8), 2387–2388 (1993)
52. C. Meyer, O. Sqalli, H. Lorenz, K. Karrai, Slip–stick step-scanner for scanning probe microscopy. Rev. Sci. Instrum. **76**(6), 063706 (2005)
53. A. Morin, New friction experiments carried out at Metz in 1831–1833. Proc. Fr. R. Acad. Sci. **4**(1), 128 (1833)
54. C.C. De Wit, J. Carrillo, A modified EW-RLS algorithm for systems with bounded disturbances. Automatica **26**(3), 599–606 (1990)
55. L.L. Liu, H.Z. Liu, Z.Y. Wu, Z.M. Wang, The application of the friction model in mechanical system. Adv. Mech. **38**(2), 201–212 (2008)
56. V. Johannes, M. Green, C. Brockley, The role of the rate of application of the tangential force in determining the static friction coefficient. Wear **24**(3), 381–385 (1973)
57. C.C. De Wit, H. Olsson, K.J. Astrom, P. Lischinsky, A new model for control of systems with friction. IEEE Trans. Autom. Control **40**(3), 419–425 (1995)
58. P.R. Dahl, Solid friction damping of spacecraft vibrations, in *1995 IEEE International Conference on AIAA Guidance and Control Conference (1995)*, pp. 1075–1104
59. P.R. Dahl, Solid friction damping of mechanical vibrations. AIAA J. **14**(12), 1675–1682 (1976)
60. D.A. Haessig, B. Friedland, On the modeling and simulation of friction. J. Dyn. Syst. Meas. Control **113**(3), 354–362 (1991)
61. N. Barahanov, R. Ortega, Necessary and sufficient conditions for passivity of the LuGre friction model. IEEE Trans. Autom. Control **45**(4), 830–832 (2000)
62. J. Swevers, F. Al-Bender, C.G. Ganseman, T. Projogo, An integrated friction model structure with improved presliding behavior for accurate friction compensation. IEEE Trans. Autom. Control **45**(4), 675–686 (2000)
63. V. Lampaert, J. Swevers, F. Al-Bender, Modification of the Leuven integrated friction model structure. IEEE Trans. Autom. Control **47**(4), 683–687 (2002)
64. K. Johanastrom, C. Canudas-De-Wit, Revisiting the LuGre friction model. IEEE Control Syst. **28**(6), 101–114 (2008)
65. J. Wang, Q. Lu, How are the behaviors of piezoelectric inertial sliders interpreted? Rev. Sci. Instrum. **83**(9), 093701 (2012)
66. A. Stieg, P. Wilkinson, J. Gimzewski, Vertical inertial sliding drive for coarse and fine approaches in scanning probe microscopy. Rev. Sci. Instrum. **78**(3), 036110 (2007)
67. K. Anantheshwara, N. Murali, M. Bobji, Effect of friction on the performance of inertial slider. Sadhana **33**(3), 221–226 (2008)
68. W.F. Smith, M. Abraham, J. Sloan, M. Switkes, Simple retrofittable long-range x–y translation system for scanned probe microscopes. Rev. Sci. Instrum. **67**(10), 3599–3604 (1996)

69. K. Svensson, Y. Jompol, H. Olin, E. Olsson, Compact design of a transmission electron microscope-scanning tunneling microscope holder with three-dimensional coarse motion. Rev. Sci. Instrum. **74**(11), 4945–4947 (2003)
70. S. Reymond, Low temperature scanning contact potentiometry. Rev. Sci. Instrum. **75**(3), 694–698 (2004)
71. J.-M. Breguet, R. Perez, A. Bergander, C. Schmitt, R. Clavel, H. Bleuler, Piezoactuators for motion control from centimeter to nanometer, in *2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2000)*, pp. 492–497
72. M. Rakotondrabe, Y. Haddab, P. Lutz, Voltage/frequency proportional control of stick–slip micropositioning systems. IEEE Trans. Control Syst. Technol. **16**(6), 1316–1322 (2008)
73. B. Sedghi, Control design of hybrid systems via dehybridization (2003)
74. H.K. Khalil, J. Grizzle, *Nonlinear Systems*, vol. 3 (Prentice Hall, Upper Saddle River, 2002)
75. J.-M. Breguet, R. Clavel, Stick and slip actuators: design, control, performances and applications, in *1998 IEEE International Symposium on Micromechatronics and Human Science (MHS)* (IEEE, 1998), pp. 89–95
76. P. Niedermann, R. Emch, P. Descouts, Simple piezoelectric translation device. Rev. Sci. Instrum. **59**(2), 368–369 (1988)
77. T. Morita, R. Yoshida, Y. Okamoto, M.K. Kurosawa, T. Higuchi, A smooth impact rotation motor using a multi-layered torsional piezoelectric actuator. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **46**(6), 1439–1445 (1999)
78. Y. Zhang, W. Zhang, J. Hesselbach, H. Kerle, Development of a two-degree-of-freedom piezoelectric rotary-linear actuator with high driving force and unlimited linear movement. Rev. Sci. Instrum. **77**(3), 035112 (2006)
79. T. Morita, R. Yoshida, Y. Okamoto, T. Higuchi, Three DOF parallel link mechanism utilizing smooth impact drive mechanism. Precis. Eng. **26**(3), 289–295 (2002)

Chapter 2
Piezoelectric Motor Technology: A Review

Amro Shafik and Ridha Ben Mrad

Abstract Piezoelectric actuators are increasingly used in various nanopositioning applications and emerging applications where miniaturization is important. This is due to their unique characteristics including their very high accuracy and short response time as compared to electromagnetic based motors and actuators and higher output force compared to electrostatic actuators. Piezoelectric motors use actuators that take advantage of the converse piezoelectric effect. In this chapter, these motors are classified into quasistatic and ultrasonic motors (USMs) based on their working frequency. Several designs from the literature and commercial suppliers are reviewed and their characteristics are presented. Two examples of piezoelectric motors are discussed in detail. These include a piezoworm stage and a USM with segmented electrodes. Future development of these technologies is also briefly discussed addressing issues such as increasing the output power, the efficiency, and further miniaturization of these devices.

2.1 Introduction

There is an ever increasing need for ultra-precision positioning technologies. These ultra-precision positioning technologies require actuators and motors with nanometer accuracy and with ranges in the order of tens of millimeters. The need is increasing in many fields including nano-metrology, biomedical and surgical devices, NEMS and MEMS, micro and nanolithography and semi-conductor manufacture, and adaptive optics [1–6]. Piezo technology is a class of actuation that is capable of fulfilling this need.

Depending on the specific application area, motors are needed to have a number of characteristics including small size, light weight, low noise levels, no electromagnetic interference [7], high accuracy and precision, short response time, high output force, simple design, and simple operation [8]. Motors and actuators can generally be divided, depending on their working principle, into six main technologies: electromagnetic, electrostatic, thermal, osmotic, electro-conjugate fluid (ECF), and

A. Shafik • R. Ben Mrad (✉)
Mechanical and Industrial Engineering, University of Toronto, Toronto, ON, Canada
e-mail: am.ibrahim@mail.utoronto.ca; rbenmrad@mie.utoronto.ca

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_2

piezoelectric actuators [8]. *Electromagnetic (EM) motors and actuators* produce forces that typically rely on the interaction between a magnetic field that is induced by a current in the stator coil and the rotor's permanent magnets. These motors have a high energy density, but unfavorable scaling laws hinder them from being miniaturized, even to the millimeter size. *Electrostatic actuators,* like electromagnetic actuators, use non-contact forces, which are produced by the interaction between materials with different charges. Electrostatic actuators have excellent scalability which makes them popular and widely used in microelectromechanical systems (MEMS) [9]. However, there are some disadvantages that limit their use, such as snapping down at the end of the output range, nonlinear output, sensitivity to the environment (e.g., humidity and ambient gas content), and the output torque or force for micro scale devices is typically limited to the pico scale. In *thermal actuators and motors,* the driving force relies on the mechanical strain of the material used. Amongst thermal actuators, shape memory alloys (SMA) have the highest strain rates. They have a relatively good scalability and high output force but suffer from relatively low response time and limited lifespan because of the plastic strain that arises with repeated cycling. *Osmotic actuators* do not require electrical input power. The driving force, in this case, is developed based on the increase of the pressure in a vessel, which in turn leads to an expansion of a diaphragm. This pressure is achieved by the osmotic effect that leads to a flow of liquid in one direction across a semi-permeable diaphragm. These actuators are suitable for microfluidics applications and are known to have a slow response time and are characterized by complex designs and different problems related to fouling and solute deposition. *ECF motors and actuators* produce force that is based on the jetting of a fluid in an electric field. The jet rotates the motor by using vanes on the rotor. This type of motors has good scalability and good output force up to sizes of the order of 1 mm^3. P*iezoelectric ultrasonic motors (USMs)* are one class of piezoelectric motors. They produce forces that rely on the converse piezoelectric effect, which converts the electric input into a mechanical strain. This type can produce useful forces in small-size actuators designed to operate at the resonant frequency of the stator. Overall, USMs offer larger torque at smaller scales and have the best potential for meeting the high performance requirements for small and micro scale devices. USMs can generate torques two orders of magnitude larger than other technologies at the micro scale [10]. Piezoelectric USMs as well as other piezoelectric motor technologies are presented in detail in this chapter. These motors use piezoelectric actuators that are typically incorporated inside mechanical structures and are repeatedly excited to generate linear or rotary motion.

Piezoelectric motors rely on integrating piezo actuators that generally offer many advantages such as large bandwidth, low thermal coefficients, high reliability, large force, unlimited resolution (sub nanometer range), no magnetic field produced, fast response (microsecond time constant), operation at cryogenic temperatures, and they come at moderate prices compared to other actuators. Other characteristics include displacements ranging from 0.001 to 0.002 of the piezo length (0.1–0.2 %), and a maximum operating temperature near 300 °C for widely used ceramics. However, they have some disadvantages including high voltage

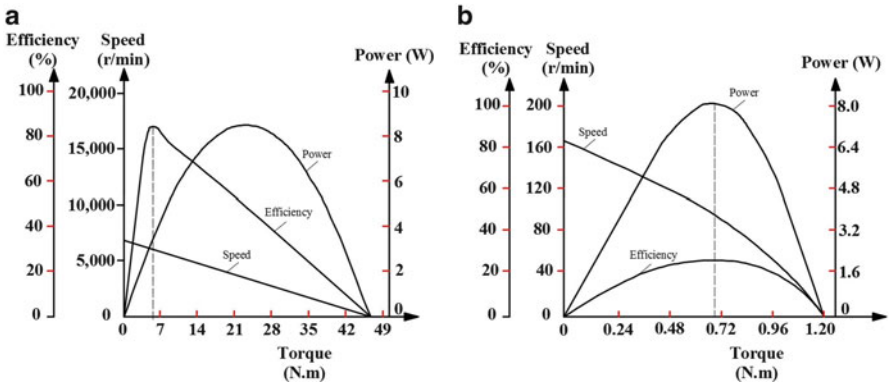


Fig. 2.1 Mechanical characteristics of (**a**) DC motors, (**b**) Ultrasonic motors

requirements (typically 100 V) increasing with the thickness of the actuator. Creep and hysteresis nonlinearities at levels ranging from 15 to 20 % are also present in these devices [11]. In addition, as all ceramics, piezoceramic actuators are brittle and require special protection and packaging. Nevertheless, piezomotors compete with magnetic motors when the requirements include small size, fast response, and high accuracy.

A comparison between USMs and DC motors characteristics was made in [7]. As shown in Fig. 2.1a, the maximum efficiency of a DC motor occurs when the output torque is minimum and the motor is operating close to the no-load speed. On the other hand, as shown in Fig. 2.1b, the maximum efficiency of a USM occurs at a higher torque and lower speed. Therefore, a USM is more suitable for operation when the operating conditions call for high torque and low speed. Another comparison between the two motors focused on their transient response characteristics is shown in Fig. 2.2a [12]. It is evident that a USM can respond faster than an electromagnetic motor with better transient characteristics. A third comparison is made with respect to the miniaturization versus the efficiency of both types of motors [13]. As shown in Fig. 2.2b, the efficiency of electromagnetic motors (EM) declines rapidly when the diameter of the motor is below 10 mm. However, the efficiency of USM changes smoothly. In addition, manufacturing of electromagnetic motors in smaller sizes is difficult because their rotors need to be surrounded by coils and their structures are more complex than that of a USM. Hence, miniaturization of USMs is more achievable than that of electromagnetic motors.

There are many other advantages of using piezoelectric USMs as compared to DC motors. The USMs have torque densities three to ten times higher than standard DC motors [8]. Therefore, smaller USMs can potentially replace DC motors and still exhibit high torque outputs. Reduction in size and weight, simple structure, and absence of gearing mechanisms lead to accurate positioning resolution, quick response, excellent controllability, and hard braking. USMs can produce high

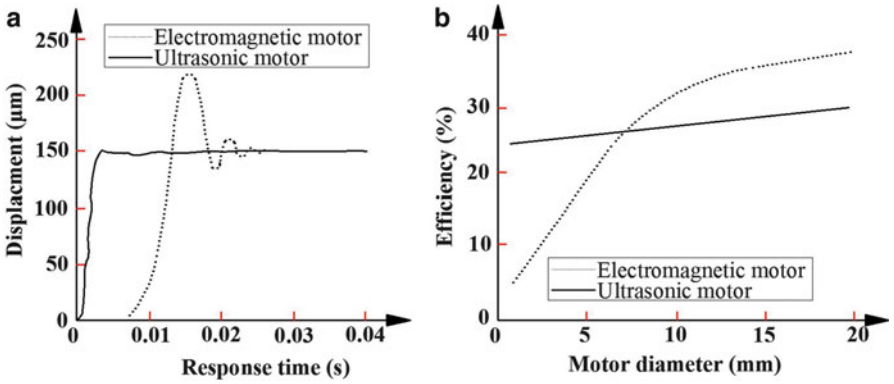


Fig. 2.2 Comparison between electromagnetic motors and USM (**a**) Basic form of transient response characteristics (**b**) Basic form of miniaturization vs. efficiency curve

holding torque even if no power is applied [14]. In addition, no magnetic or radioactive fields are generated in USMs, and their effects are negligible on them. The use of ultrasonic vibration leads to silent operation. The applications for USMs include autofocus systems in cameras [7], micro-robots [8, 10], and micro-actuators for catheters or tele-surgery operations [15].

Piezoelectricity was discovered by the Currie brothers in 1880 in materials like Quartz and Rochelle salt. In the 1940s, the discovery of barium titanate ($BaTiO_3$) led to the beginning of an industrial era for piezoelectric actuator technology. Early piezoelectric actuators and motors were mainly used in the 1960s in the watch industry. In 1965, Lavrinenko developed the first ultrasonic piezomotor [16]. In the 1970s, Siemens developed a step piezomotor [7], and in 1973, IBM developed an ultrasonic piezomotor based on longitudinal vibrations of horns [17]. In 1980, a Langevin vibrator was first used to excite the longitudinal vibration mode in USMs [18]. A significant advance in piezomotor technology occurred in 1982 and 1983 when Sashida developed the standing wave and traveling wave USMs, respectively [19, 20]. In 1985, Kumada developed a USM based on a hybrid type of longitudinal and torsional vibrations that was driven by a single-phase signal [21]. In 1987, Panasonic developed a ring-type traveling wave ultrasonic piezomotor using Sashida's design [22]. In 1987, Shinsei© [7] was the first company to commercialize USMs.

In this chapter, a detailed review of piezoelectric motor technology is presented. The chapter is organized as follows. The basic working principles of various piezomotors are discussed in Sect. 2.2. A comprehensive classification of the piezomotor family is presented in Sect. 2.3. Several linear and rotary designs of piezomotors are reviewed for quasistatic and ultrasonic types in Sects. 2.4 and 2.5, respectively. Examples of piezomotors are presented in Sect. 2.6, and future developments are listed in Sect. 2.7.

2.2 The Piezoelectric Effects

Piezoelectric materials are known to have two effects. The direct piezoelectric effect is the ability of the piezoelectric material to generate an electric charge when it is subject to a mechanical stress. The converse piezoelectric effect is the ability of the material to generate stress when it is subject to an electric field. The direct effect is utilized mainly in sensors, while the converse effect is utilized in actuators. The working principle of piezomotors is based on the converse piezoelectric effect which is based on having one or more piezoelectric actuators integrated into a mechanical structure. These piezoelectric actuators are powered leading to generating a strain in the piezo material.

The output strain in piezoelectric actuators is very limited, about 0.1–0.2 % of the length of these actuators. Therefore, in various applications and in order to achieve a useful strain, the initial strain is amplified by using amplifying mechanisms or by simply repeating the small steps of the piezoelectric actuator at a high frequency leading to macroscopic motion. This latter approach is typically the case with USMs where the piezoelectric motors are excited repeatedly at one of their resonant frequencies. Compliant structures are often used as amplifying mechanisms and ways to channel the strain to the load especially when the motors are operating at high frequencies and this leads to high precision due to the absence of backlash and ease in miniaturization since typically there are no moving parts incorporated into the structure of the motor except for the rotor or moving load.

2.3 Classification of Piezomotors

Piezomotors are classified in the literature according to a number of criteria including the wave propagation method, the output motion and its direction, the geometric shape of the stator, and the vibration modes used and their combinations [8, 14, 15, 23, 24]. The most common classification is based on the driving wave generation method. Zhao [7] added additional features such as the contact status between the stator and rotor, the nature of the excitation of the stator by the piezoelectric actuators (resonant versus non-resonant), the direction of the operating mode (in-plane versus out-plane), and the number of degrees of freedom of the rotor.

Piezomotors are classified in this report according to the motion type into linear and rotary motors, and according to the input frequency into quasistatic, in which the input frequency is relatively low and does not correspond to any of the resonant frequencies of the motor, and ultrasonic, in which the frequency is above 20 kHz and the excitation frequency often corresponds to a resonant frequency of the motor. Quasistatic motors, in turn, are classified into inertial and stepping types based on their principle of operation. The USM family is subdivided into a number of subfamilies. Contact and non-contact types USMs. In the contact type, the stator and rotor are in contact by means of a friction material that transfers the vibration

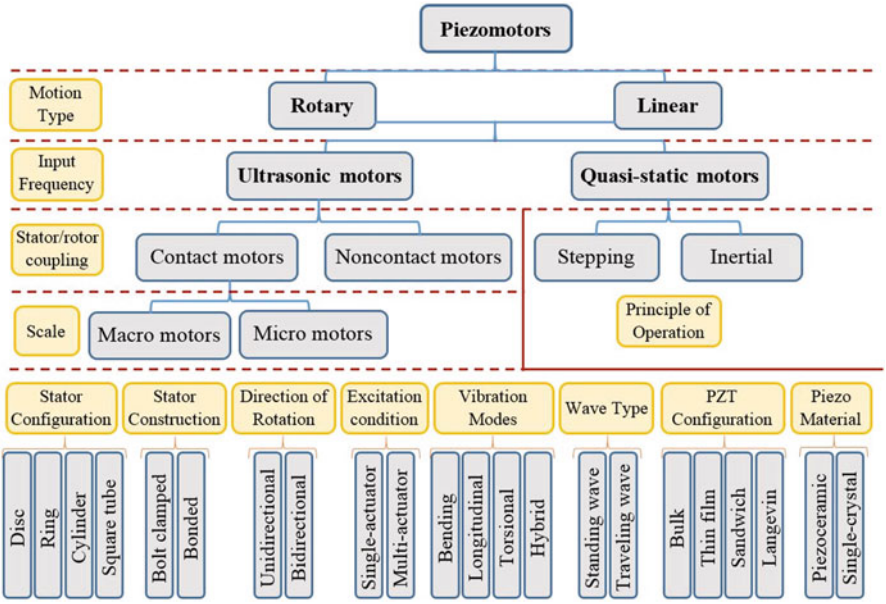


Fig. 2.3 Taxonomy of piezomotors

of the stator to the rotor. The non-contact type uses a fluidic medium instead of a friction material. USMs can be also categorized into micro motors, in which the size is few millimeters and smaller, and macro motors, which are larger than three millimeters. USMs can be labeled based on the stator shape and construction, direction of rotation, number of driving piezo actuators, type of exciting vibration signal, type of generated driving wave, form of piezo elements, and piezo material used. This entire taxonomy is shown in Fig. 2.3.

2.4 Quasistatic Motors

Quasistatic motors are a type of piezomotors that work with an input voltage at a frequency lower than any of the resonant frequencies of the motor. Quasistatic motors are classified into stepping and inertial types based on their principle of operation. Most of the quasistatic motors are of the stepping type. In this section, both types are discussed for both linear and rotary type configurations.

In the stepping types, the motor consists of several piezo actuators that are integrated within a mechanical structure in a way to generate motion by a sequence of expand–contract and/or clamp–unclamp cycles. In the inertial types, the motor works mainly based on the inertia of a moving mass which is connected to a metal shaft which in turn is attached to a piezo actuator. Generally, quasistatic motors have higher resolution and force but lower speed than USMs.

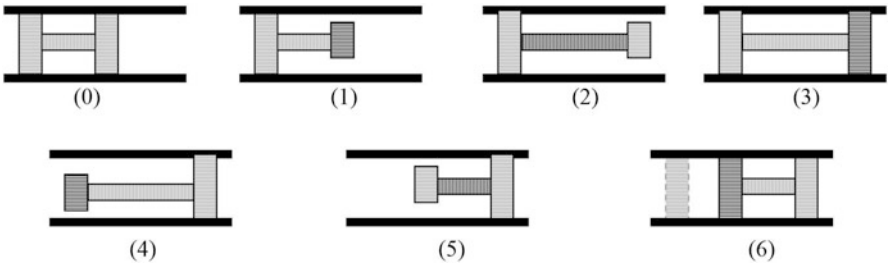


Fig. 2.4 Inchworm stepping sequence

2.4.1 Linear Quasistatic Motors

Inchworm Motors The most common and simplest type of quasistatic motors is the inchworm motor [25–27]. These motors consist typically of three piezoelectric actuators; two are used for clamping and one for extension to generate the desired motion of the load. The inchworm uses a sequential clamping mechanism to inch to the required position (see Fig. 2.4). Several designs based on the same principle with a number of variations have been proposed [28–30]. One of them is the caterpillar motor [31] in which five piezoelectric actuators are used instead of three to provide extra clamping and extension in order to have a continuous motion (i.e., smoother motion). The caterpillar motor is more complex than the three-actuator based inchworm. Inchworm motors are typically pursued when nanometers accuracy is required [32].

Walking Motors As the name implies, the walking motor uses two or more sets of piezoelectric actuators for enabling the motion steps or walks. The arrangement of actuators makes one actuator or set of actuators used for clamping while the other set is used to move the rail [33]. In the PiezoWalk Drive®, the system is composed of two sets of piezoelectric actuators and two legs for pushing the moving element. Each set consists of two multilayer piezo stack actuators. One set is responsible for pushing the legs to enable motion, and the other set is used for engaging and disengaging the legs. One step of the drive piezos moves the runner up to 15 μm. Subsequently, the other clamp engages with the other leg to enable a continuous motion [34].

The PiezoWalk® (produced by Physik Instrumente [35]) and the Piezo LEGS(R) (produced by Piezomotor AB [36]) work based on the same stepping principle but they use different arrangements. In the PiezoWalk® motor [35], the clamping and feeding actions are achieved by longitudinal and shear actuators contained in four legs. The legs are working in pairs for the clamping and the feeding motion. In the Piezo LEGS®, the motor also has four legs, but each of them is a piezoelectric bimorph actuator [37]. The principle of operation is as follows: at the beginning of the cycle, all legs are elongated and thus bend. Then, the first pair of legs moves

towards the right while maintaining contact with the runner while the second pair retracts and moves to left. The second pair then extends to contact the runner and pushes it to the right while the first pair retracts and moves to the left. And the cycle continues [36].

Inertial Motors Inertial motors are composed of a single piezo actuator, a moving mass, and a drive shaft. This type is classified as a quasistatic motor [24], even though its operating frequency can reach up to 20 kHz. The basic idea of the inertial piezomotors is that when the piezo actuator expands and contracts at a low acceleration, the moving mass and the shaft move together because of friction, and when the piezo expands and contracts at a high acceleration, the inertia of the moving mass makes it slide on the shaft. The low acceleration mode is achieved by increasing the input voltage slowly while the high acceleration mode is accomplished by decreasing the voltage rapidly. This can be done if the piezo actuator is controlled with a modified sawtooth voltage waveform. Reversing the motion is easily obtained if the waveform is mirrored. An example of an inertial-type piezomotor is the one used in the anti-shaking mechanism in cameras made by Konica Minolta®. The PIShift® inertia drive [35] is an example of a commercial inertial piezomotor. The actuator expands slowly taking along the runner, but when it contracts rapidly, the runner cannot follow due to its inertia, as shown in Fig. 2.5. The main drawback of this motor is the high wear rate that is caused by the sliding friction.

Typical quasistatic motors incorporate several piezoelectric actuators except the inertial piezomotors. Quasistatic motors are better than USMs in terms of control, because both the amplitude of the driving signal and the frequency can be controlled to achieve a desired performance. These motors also offer higher resolution and

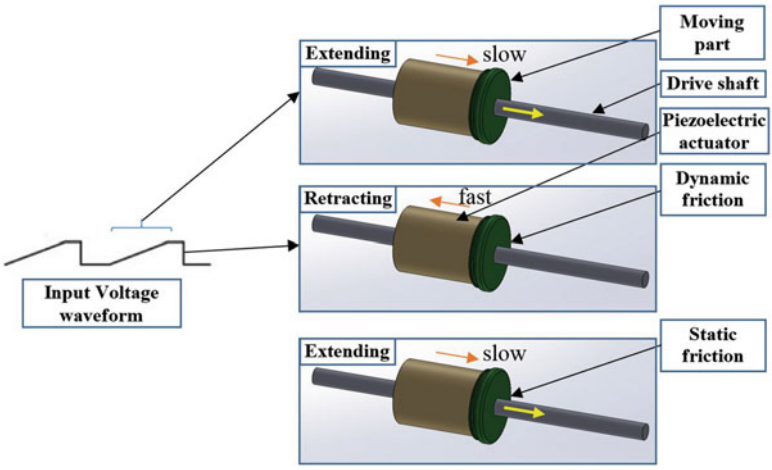


Fig. 2.5 Inertial piezomotors (PIShift® operating principle)

Table 2.1 Commercial linear quasistatic piezomotors and some of their key specifications

Product	Manufacturer	Concept	Force (N)	Speed (mm/s)	Accuracy (nm)	Range (mm)
Inchworm®	EXFO Burleigh	Stepping	10	1.5	4	50
PiezoWalk®	Physik Instrumente	Stepping	10	20	20	30
Piezo LEGS(r)®	PiezoMotor AB	Stepping	450	0.3	<1	20
PIShift®	Physik Instrumente	Inertial	10	10	<1	26

force than USMs. Increasing the number of actuators integrated into the motor can make the motion smoother but this comes at an increased complexity and cost. A list of commercial quasistatic motors is shown in Table 2.1.

2.4.2 Rotary Quasistatic Motors

Inchworm Motors The development on this type of motors has not been pursued as much as that of linear inchworm motors. In the case of rotary motors, most of the attention has been directed towards rotary USMs. The rotary inchworm motors are well suited as rotary positioners in applications requiring high accuracy and resolution.

The operating principle for linear and rotary inchworms is the same. The rotary type has typically two sets of clamps and a mechanism to rotate the motor by stepping action. One clamp is fixed to the stator (the stationary part of the motor) and the other clamp is attached to the rotor (the rotating mechanism). Rotary inchworm piezomotors can be grouped into three types; *disc motors*—clamp on face of the disc [38, 39], *drum motors*—clamp on inside diameter of drum [40], and *shaft motors*—clamp on outside diameter of shaft [41, 42].

Various approaches have been used for the clamping and rotation mechanisms. The rotation mechanism can be achieved by using a torsional piezo actuator [39] or by using a linear actuator and a flexure to convert the output linear motion into rotational motion [38, 40, 41]. The problem with using a torsional actuator is the likelihood of failure of the piezo material since it is subject to a direct shear load. The use of flexures protects the piezo and provides a constant preload that keeps it in compression, but this comes at the expense of an increased size and complexity of the motor [43–45]. Regarding the clamping mechanisms, the clamps used in the disc and drum motors are flexures driven by linear actuators [38, 40, 41]. In shaft motors, the clamps are in the form of tubular piezo elements that directly clamp the shaft. A drawback of this latter type of clamps is the high wear and the likelihood of failure due to shear forces [42]. Thus, the use of flexures with linear actuators is preferred for clamping and rotation mechanisms [25].

The selection of the inchworm type to use depends on the application requirements. By comparing the three types of rotary inchworms listed above leads to the

following: The shaft motors have a better response time than the disc and drum motors because of the lower inertia and their relatively small size. The shaft motors though provide a smaller torque than the other types. This is so because the torque is directly proportional to the clamping radius. The clamping radius being the distance from the axis of rotation to the clamp point. They also have a large step size since the step size is inversely proportional to the clamping radius. This implies that shaft motors can reach higher speeds but at the expense of resolution. Therefore, for applications requiring high torque and resolution, the disc or drum motors are the better choices. If speed is the main criterion, then the shaft motor is a better option.

Stepping Motors Rotary stepping motors work based on the same principle as the linear ones. PiezoMotor AB offers a rotary Piezo LEGS motor. The smallest of these motors has a 17 mm diameter, a stall torque of 30 mNm, an angular speed of up to 28 rev/min, and an angular resolution of 0.1 μrad.

2.5 Ultrasonic Motors

USMs are piezomotors that operate based on an input voltage in the ultrasonic frequency range (usually above 20 kHz) and with the input signal frequency corresponding to one or more of the resonant frequencies of the motor. USMs come in different sizes (i.e., micro versus macro) and can be classified into contact and non-contact motors. In this section, all types are discussed for both linear and rotary motion.

2.5.1 Working Principle

The theory of operation of piezoelectric USMs is based on the converse piezoelectric effect. This concept is applied to convert the input voltage to the piezoelectric ceramic stator into resonant vibrations at a high frequency in the range of 20 kHz–10 MHz, leading to a linear or rotary motion transmitted to the rotor often through friction between the rotor and stator [24]. Thus, the working principle of USMs consists of a two-step process:

1. *Electrical energy* is converted into *mechanical vibrations* in the stator, which includes the piezoelectric elements, leading to mechanical vibrations at ultrasonic frequencies.
2. At the stator–rotor interface, the high frequency *vibrations* of the stator are transformed into macroscopic rotor *motion* (translation or rotation) by means of frictional forces between stator and rotor.

A functional block diagram for rotary USM is shown in Fig. 2.6.

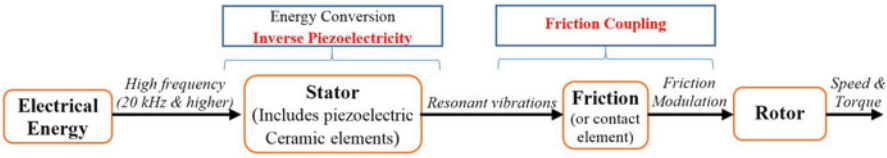


Fig. 2.6 Functional block diagram for USMs

A key aspect in producing the motion of the motor from the vibrating stator is how to generate a stator tip motion and subsequently how to transmit it to the rotor. The friction coupling that often incorporates an initial preload force between stator and rotor ensures that the stator and rotor are constantly in contact and that the output is repeatable every stator cycle. The preload is a normal force applied to the rotor towards the stator. Because of the simplicity of USMs, a widely used motor control method based is pulse width modulation (PWM) [15].

2.5.2 Linear Ultrasonic Motors

A number of designs have been presented by various research groups and industry. The most prominent of which are listed below.

Elliptical Push Motor The elliptical push motor is a single actuator USM that uses geometric coupling of two eigenmodes; the longitudinal vibration mode and the bending mode (1L2B), to produce elliptical motion at the drive tip of the piezoelectric actuator. This arrangement has been used in the Paderborn rowing motor [46] and commercialized by Nanomotion [47] in their HR series. This motor consists of a rectangular piezoelectric plate and four electrodes. It works by exciting the motor at the resonant frequencies of the piezo plate. The advantage of this design is that it eliminates the need for frequency coupling of the two modes and it uses only one amplifier. However, its main problem is that it is difficult to increase the mechanical power developed by the actuator by increasing the size of the piezoelectric plate because the power is directly proportional to the square of the frequency, and the resonant frequencies decrease as the size of the piezo increases [48].

Traveling Wave Motor Traveling wave motors, also known as propagating wave motors, work by exciting piezoceramics that are attached to the stator to generate two resonant standing waves with a phase shift of 90°. This results in an elliptical motion at the surface of the stator, which causes the rotor to move with a speed determined by the horizontal component of the speed [49]. Traveling wave motors are usually configured for rotary motor applications (to be discussed in Sect. 2.5.3) as a circular path is needed for the wave to travel. Examples for linear traveling wave motors are also available such as the ring and vibrator motors that were developed

by Sashida [49] in 1982. The main drawback of this motor is the low efficiency (theoretically less than 50 % due to the requirement of two drive signals to generate the traveling wave).

π-Shaped Motor Uchino and Ohnishi [33] presented the π-shaped motor. The stator has fork-shaped metallic legs and is driven by either one or two multilayer piezoelectric actuators. A slight difference in the bending mode resonant frequency of the two legs creates a phase shift between them and allows the motor to move the rail with a walking motion like a trotting horse. The motor design has two actuators at two corners of the π-shaped frame and is driven by a propagating wave. This is achieved by driving the two piezo actuators with a 90° phase difference.

E_{31} Linear Motor An elliptical motion at the drive tip is used to drive a load and can be achieved by different concepts and arrangements. Elliptical push motors (discussed previously) use the 1L2B vibration mode, which is accomplished by exciting simultaneously two different eigenmodes (longitudinal and bending) on a piezoelectric plate. Physik Instrumente utilizes the E(3,1), or simply E_{31}, vibration mode to a single actuator to create an elliptical motion at the drive tip. The E_{31} vibration mode is a two-dimensional standing extensional wave mode in which the *x*-component is at the third vibration mode and the *y*-component is at the first vibration mode. To achieve this vibration mode, the dimensions of the piezo plate should be such that the *X*:*Z* (length-to-height) ratio is around 2:1. The motor consists only of a piezo plate with two electrodes at the front side and one common electrode at the back. The motor works in one direction by using one of the front electrodes to apply a potential across the piezo and changes its direction by using only the other electrode to apply the potential. Some of its advantages include simplicity, the highest efficiency amongst linear piezomotors, and scalability [48]. The current specifications of the PILine® ultrasonic piezomotors [35] are: velocity up to 500 mm/s, holding force up to 15 N, and resolution of 2 nm.

Nanomotor The Nanomotor [50] is a concept developed by Klocke Nanotechnik [51] which uses one piezo tube. For fine positioning, the extension is carefully controlled by the piezo tube, but for coarse positioning the piezo is excited with an ultrasonic pulse wave. The main specifications include a speed up to 5 mm/s, driving force of 0.25 N, 10 mm positioning stroke, 20 nm/V fine positioning stroke, and atomic resolution.

2.5.3 Rotary Ultrasonic Motors

Rotary USMs can be classified according to the type of the wave that is used to excite the piezoelectric material into two types, standing wave type [52, 53] and traveling (or propagating) wave type [54–56]. Traveling wave motors are the most common type [57]. The traveling waves are generated by combining two standing

waves at the same frequency and amplitude but with a phase difference of 90°. This leads to the formation of an elliptical motion at the interface of the stator and the rotor. A comparison of the two types of USMs is shown in Table 2.2.

Another classification of USMs can be done according to their assembly method, or the operating mode of the piezoceramic material, commonly using PZT (lead zirconate titanate) as in [58]. Motors can be classified into bonded type [59] and bolt-clamped type [52, 56]. The latter type gives higher efficiency and output power than the former one. This is due to the adoption of the d_{33} mode of PZT in bolt-clamped type motors, which have higher transfer efficiency than the d_{31} mode typically used in bonded type motors [56, 60]. In the bonded type, the stator is constructed by gluing PZT ceramic to elastic metal blocks with conductive adhesive. This construction leads to low torque and low efficiency [61] due largely to the lower electromechanical coupling factor of the d_{31} mode of PZT, the PZT and the adhesive layer fatigue under the intense vibration, and the difference in the vibration amplitudes of the axial bending mode in axial and circumferential directions. Bolt-clamped [52] or sandwich-type transducers [57, 61] are thus preferred if the above listed issues are of concern. However, bolt-clamped motors have issues related to the high energy consumption. In a number of designs, Langevin-type transducers [54, 59] have been adopted because of their excellent energy efficiency, stability, large output torque, high reliability, and good environment adaptability [58]. Langevin transducers [62, 63] are piezo-based transducers that convert electrical energy to ultrasonic mechanical vibrations with amplified amplitude. The use of Langevin transducers in USMs resulted in better characteristics and opened new application areas including aeronautics and astronautics [59]. A review of different macro-USMs is shown in Table 2.3.

A number of rotary designs have been proposed. They include the following.

Wedge and Traveling Wave The wedge and the traveling wave types are two of the most prominent designs of rotary USMs designed by Sashida [4]. In the wedge type motor, a Langevin vibrator is used to produce an amplified ultrasonic wave, which is used to create elliptical motion at the tip of one or more wedges. The wedge is pushed against a disc to convert the tip elliptical motion to a continuous motion. This motor could achieve speeds up to 3000 rpm but suffers from rapid wear. The traveling wave motor does not have the same wear issue. The operation is the same as for the linear traveling wave motor which was described in Sect. 2.5.2. This motor is commercially being used in the auto focus system of the Canon cameras. It is more compact than traditional motors, has a fast response time, and silent operation.

Twist Connector The twist connector motor uses two vibration modes, longitudinal and torsional. It is called also the mode conversion type motor [23] because the stator converts the longitudinal vibration that is produced by the piezoelectric disks to a torsional vibration by an elastic torsion coupler. While this motor is more complex, it has the advantage of higher power density.

Elliptical Push This type of motor uses the same principle as the linear elliptical push motors except that the slider is circular instead of being linear [34]. Anorad

Table 2.2 Comparison between standing and traveling wave piezo ultrasonic motors

Standing wave	Traveling wave
Advantages	
– An efficiency of up to 98 % theoretically [15] and 87 % achieved practically [10]. – Lower cost because of using one vibration source only. – Easier to support the stator at the nodal points or lines. This causes minimum effects on the resonance vibration [15].	– Less abrasion on the contact surfaces because the traveling wave drives the rotor continuously [7]. – The high frequency operation allows for a very high accuracy to be obtained from the rotor or slider (sub-nanometer positioning accuracy) [23]. – Controllable in both rotational directions [15] because the motor is a dual-signal drive, that means it consists of two sets of transducers that are excited independently.
Disadvantages	
– The standing wave drives the rotor discontinuously, which causes wear on the contact surfaces [7]. This may lead to a short lifespan due to the repeated stator–rotor collision [10]. – It provides intermittent rotational torque or thrust. However, because of the inertia of the rotor, the rotation speed ripple is not large [15]. – Lack of control in both directions in many designs [15, 52]. However, some designs already solved this problem, as in [53].	– Used in applications where the motor is activated for short time durations as heat dissipation is challenging [23]. – Miniaturization to sub-millimeter scales is difficult due to the difficulty to fabricate the complex geometry of the stator, dual-signal drive, and alternating poling pattern required by the transducer [10]. – The amplitude of the traveling wave has a detrimental effect on the scalability of the design [8], because the amplitude of the wave needed reduces linearly with the size of the device [8]. – More expensive because two vibration sources are required to generate one wave, and this leads to lower efficiency (not more than 50%) [15].

Table 2.3 Macro-USMs construction and published characteristics (when information is not available, the box is left blank)

Ref	Construction			Vibration mode	Type of excitation	Material	Input voltage	Frequency (kHz)	Speed (rpm)	Torque
	Stator geometry	Assembly	Transducer							
[52]	Cylindrical	Bolt-clamped	Sandwich (1)	Bending	Standing wave	PZT-41	200 V_{rms}	28.1	165	0.45 Nm
[53]	Square-tubular	Bolt-clamped	PZT square tube		Standing wave		50 V	46	1000	370 μNm
[54]	Ring	Bolt-clamped	Langevin (4)	Bending	Traveling wave	PZT-8H	250 V_{0p}	39	120	1.6 Nm
[55]	Cylindrical	Bolt-clamped	Sandwich (4)	Longitudinal	Traveling wave	PZT-41	200 V_{rms}	35.35	290	1.95 Nm
[56]	Cylindrical	Bolt-clamped	Sandwich (1) composite	Longitudinal & Bending	Traveling wave	PZT-41	200 V_{rms}	31.14	281	1.2 Nm
[57]	Square driving feet	Bolt-clamped	Sandwich—Driving feet (4)	Bending	Traveling wave	PZT-41	200 V_{rms}	25.1	298	9.3 Nm
[58]	Square driving feet	Bolt-clamped	Sandwich—Driving feet (4)	Longitudinal		PZT-41	200 V_{rms}	23.1	71	12.3 Nm
[59]	Ring	Bolt-clamped	Langevin (4)	Bending—in-plane		PZT-8H	200 V_{0p}	48.6	100	0.3 Nm
[60]	Square driving feet	Bolt-clamped	Sandwich—Driving feet (4)	Bending		PZT-41	200 V_{rms}	20	58	9.5 Nm

(continued)

Table 2.3 (continued)

Ref	Construction			Vibration mode	Type of excitation	Material	Input voltage	Frequency (kHz)	Speed (rpm)	Torque
	Stator geometry	Assembly	Transducer							
[61]	Ring	Bolt-clamped	Nested stacks	Bending (radial)	Traveling wave	PZT-41	150 V	19.2	126	0.8 Nm
[74]	Tubular	Bolt-clamped	PZT tube	Bending		$d_{31} = 15e-12$ C/N	50 V_{pp}	27.6	400	300 μNm
[75]	Bars		V-Type	Longitudinal		TOKIN—NEPEC 6	130 V_{rms}	43	700	0.75 Ncm
[76]	Ring	Bonded			Traveling wave	PZT-Pb(Zr/Ti)O_3 (C-213 model)	100 V_{0p}	Needle Bearing	60	4.2×10^{-4} N.m
[77]	Cross-shaped (plates)	Bonded	Sandwich (8)	Bending			24 V	80.7	90	180×10^{-4} N.m
[78]	Disk	PZT on SOI			Traveling wave	Thin film PZT Zr/Ti ratio of 52/48 (thickness 1000 nm)	10 V	284	2300	
[79]	Ring	Bolt-clamped	Langevin (4)	5th in-plane flexural	Traveling wave	PZT-8H	200 V_{0p}	48.65	80	0.35 Nm
[70]	Ring	Bonded	Two-sided PZT	vibration modal (6,0)	Traveling wave	PZT-8	150 V_{pp}	21.95	76	0.5 Nm

Table 2.4 Commercial rotary ultrasonic piezomotors and their specifications

Product	Manufacturer	Concept	Torque (mNm)	Speed (rpm)	Accuracy	Weight (g)
PILine®	Physik Instrumente	Standing	50	180	35 μrad	300
USR30	Shinsei Motors	Traveling	50	300	–	17

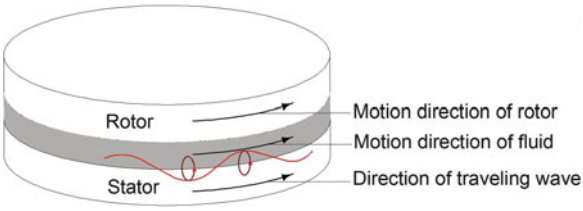


Fig. 2.7 Driving mechanism of non-contact USMs

commercialized a motor based on this concept and claims an accuracy of 3 arcsec, a significant improvement over conventional technologies.

Specifications of two commercial rotary USM produced by two leading companies in the field, Physik Instrumente and Shinsei Motors, are shown in Table 2.4.

Miniaturizing USMs to milli/micro scale is an ongoing field of research [8, 10, 64, 65]. There are two development directions in the field of micro USMs: 1. Using small bulk piezo elements [64], or 2. Using piezoelectric thin film technology [43]. In [64, 65], bulk Lead-Zirconate-Titinate (PZT) [64] was compared to a stator transducer fabricated using PZT thin film [65]. The results showed that the bulk PZT transducer gives more than thirty times larger output torque. The piezoelectric factor e_{31} was used to compare both designs as they used a tube transducer. The output torque is directly proportional to the e_{31} value. The e_{31} piezoelectric constant had a higher value in PZT bulk transducers (-4.1 C/m^2) as compared to that of PZT thin film (-0.19 C/m^2). The bulk PZT based ultrasonic micro motor provides a max speed of 650 rpm and a maximum torque of 0.22 mNm at a 100 $V_{p\text{-}p}$ input voltage and a 85 kHz driving frequency. The maximum torque produced by the PZT thin film based motor [65] is only 7 μNm.

Non-contact USMs form another research area in the USMs field. Contact USMs use the friction between the stator and rotor to transfer the motion from one to the other which leads to heat generation, lower efficiencies, and shorter life span. When high rotational speeds are needed, the use of solid–fluid–solid driving mechanisms instead of the solid–solid interface used in contact type motors was investigated [7]. When using the solid–fluid–solid mechanism, the torque is transferred to the rotor through a fluid contained in the gap between the stator and rotor leading to the rotor rotation in the direction of the traveling wave as shown in Fig. 2.7. The main features of the non-contact USMs are: there are no friction losses as in contact type motors, longer life spans, higher velocities can be achieved, and smaller torques are produced. These features make this type of USMs suitable for applications that need higher velocity and smaller torque as shown in Fig. 2.8.

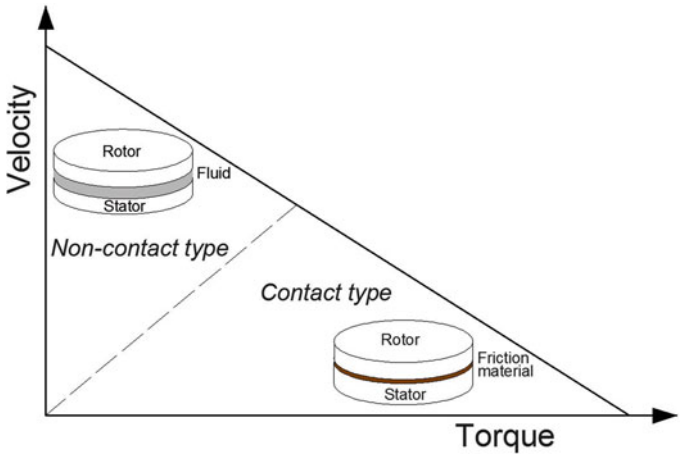


Fig. 2.8 Contact versus non-contact USMs characteristics

Increasing the stator vibration velocity and utilizing the axial resonance of the gap between the rotor and the stator can lead to an increase in the revolution speed [66]. Other features can also lead to an enhanced performance such as adding an acoustic reflector over the rotor can increase the maximum speed 1.4 times because of the existence of two air gaps, the gap between the stator and rotor and the gap between the rotor and the reflector. By using a double-disk rotor with a small air gap (0.8 mm) between the rotors, in addition to a reflector, can increase the speed by a 1.6 factor [67]. By using this latter arrangement, a speed of 2000 rpm was achieved with a very low driving voltage of 3 V [68]. The non-contact type of USMs have not been commercialized yet. A summary of the characteristics of a number of non-contact USMs is shown in Table 2.5.

2.6 Case Studies

In this section, we present two case studies as a part of research conducted at the Mechatronics and Microsystems Design Laboratory at the University of Toronto. The first case presents an Inchworm (called "Piezoworm") stage design [32], as an example of a quasistatic piezomotor. The second case presents a Linear E_{31} motor with segmented electrodes [69], as an example of an ultrasonic piezomotor.

2.6.1 Piezoworm Stage

A compact clamp piezoworm stage is integrated into a two-axis positioning system to be used for tracking profiles with nanometer accuracy. To avoid the backlash

Table 2.5 Summary of non-contact USMs designs and their characteristics (when information is not available, the box is left blank)

Ref	Gap (mm)	Speed (rpm)	Torque (N.m)	Medium	Traveling wave mode	Size (mm)	Driving frequency (kHz)	Input voltage
[7]		3000				20×1.7		
[7]		1200				30×0.5		
[7]		6031	3.5×10^{-5}			45×1.5	18.9	
[67]	0.25	2800		Air		30 (dia.) $\times$ 0.4 (th.)	25.9	
[80]	0.05	3200		Air	2 Langevin transducers			
[81]	0.7	3000			B_{03} mode			
[82]		4000			Two B_{03} modes	5 (dia.) $\times$ 0.5 (th.)	20.5	30 V_{p-p}
[83]		2026	184×10^{-6}		B_{03} mode		19.95	
[84]		5300	3.5 mg.m					
[85]		2100	1.3×10^{-5}					
[86]		3569			B_{22} mode		45.6	20 V

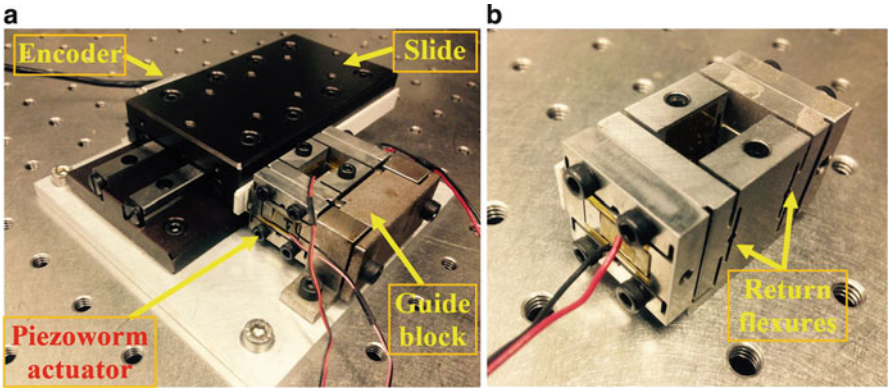


Fig. 2.9 Piezoworm design (**a**) Stage prototype along one axis, (**b**) Piezoworm motor

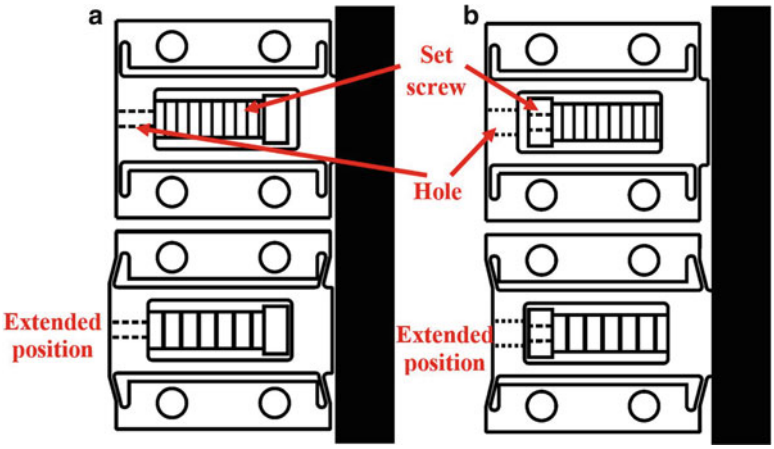


Fig. 2.10 Clamp configurations (**a**) NC clamp, (**b**) NU clamp

problem and for high precision alignment, a direct connection is designed to connect the stage to a commercial slide. Prototypes of the piezoworm stage and motor are shown in Fig. 2.9.

The piezoworm stage consists of a piezoworm motor, a crossed-roller slide, and a 10 nm resolution encoder, all mounted on a small square base plate of 120×120 mm dimensions [32]. The piezoworm actuator dimensions are $27 \times 43 \times 25$ mm. It has two clamps, the normally clamped (NC) clamp, which grips at zero voltage, and the normally unclamped (NU) clamp, which grips when voltage is applied. The clamps are connected to a middle section by extension flexures which work as return springs to preload and protect the piezo stacks. The piezo stacks for both clamps are made of PZT Navy type II which offers good free expansion and high stiffness. Friction surfaces are covered by ceramic strips to reduce wear. The clamp configurations are shown in Fig. 2.10. The flexures are almost the same for both clamps and the only

difference is the bottom hole, which is a tapped hole for a set screw in the NC clamp, and a through hole in the NU clamp. This has been made for adjusting the set screw to preload the piezo stacks. When piezo stacks are energized, they expand and push the clamping surface towards the slide in the NU clamp, and away from the slide in the NC clamp.

The piezoworm stage with the closed loop control showed its efficacy with average positioning accuracy of ± 20 nm and an average error in tracking applications when connected to a two-axis stage of 8 nm, 1.72 μm, and 1.85 μm, in the nanometer, micrometer, and millimeter regimes, respectively [32].

2.6.2 *Linear E_{31} Motor with Segmented Electrodes*

This motor was developed in order to solve the dead zone problem of ultrasonic piezomotors and to design a linear USM to integrate into a meso-Milling machine tool. The dead zone problem is that the output force and speed of USMs are coupled and the only control parameter is the amplitude of the input voltage.

The motor has been designed based on the E_{31} vibration mode. E_{31} is a planar standing wave mode and 3 and 1 represent the number of half-wave lengths along the *x*-axis and the *y*-axis, respectively. Figure 2.11 shows that the maximum displacement occurs at the drive tip location.

Stator dimensions were selected based on the aspect ratio of the E_{31} mode requirements to be $60 \times 30 \times 9$ mm. The material used is PZT (Navy Type II), and supports are added as shown in Fig. 2.12 [69].

Electrodes were arranged to apply a new segmented electrode concept [69]. The back surface of the motor has one common drain electrode. The front surface has multiple electrodes with different dimensions (as shown in Fig. 2.13). Based on the fact that when the active electrode length changes, the output speed and force change accordingly, the motor was tested with different active electrode lengths to decouple the output force and speed. Figure 2.13 shows simulated performance curves showing speed versus electrode size and force versus electrode size. The figure clearly shows a region where low speed and high force are feasible.

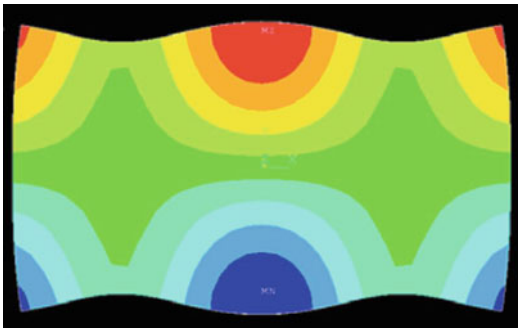


Fig. 2.11 Simulation of vertical displacement of the E_{31} vibration mode by ANSYS FE Software

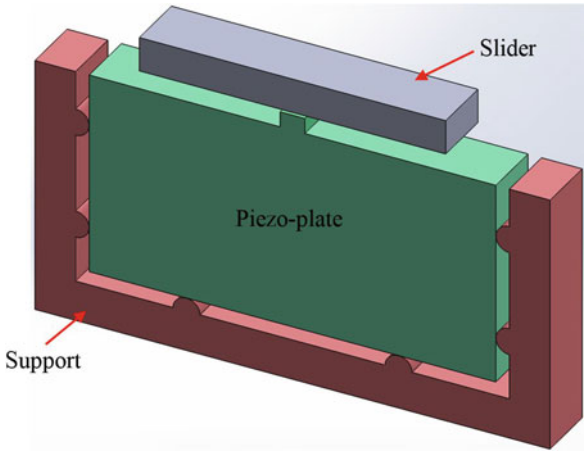


Fig. 2.12 Model of the piezo plate with drive-tip and simplified support structure and slider

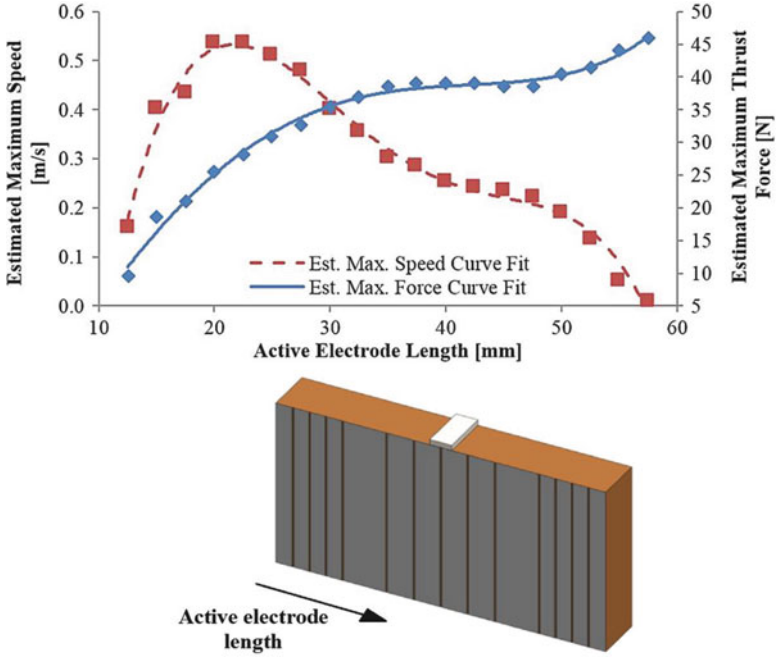


Fig. 2.13 Simulated USM performance vs. active electrode length

The experimental performance characteristics of the motor with 200 Vp and at a 70 N preload are a maximum speed of 0.223 m/s and a maximum thrust of 36 N. The smallest achievable step in open-loop was 9 nm with an 18 μs impulse. Moreover, the proposed design allowed the speed to vary while keeping the thrust force relatively constant.

2.7 Future Developments

Increasing the output power, the efficiency, and miniaturization of USMs continue to be the main areas of current research in this field. Different ideas have been investigated for increasing the output torque. Li et al. [70] improved the torque and torque density by sandwiching the stator with two piezoceramic layers bonded to its upper and lower surfaces and by adding two disc rotors. The stator has teeth on both sides. This design is capable of producing two times the torque generated by conventional USMs (i.e., with only one piezoceramic) using the same piezoelectric material and dimensions. Lu et al. [71] proposed using shear-type piezoceramics and utilizing the d_{15} piezoelectric mode for vibration excitation of the stator instead of the d_{31} and d_{33} modes because of its higher values (as shown in Table 2.6). However, the fabrication process of piezoceramics operating in the d_{15} mode is more complex because the electrode surfaces that are used for polarization are different from those used for excitation.

$$\eta = \frac{P_{\text{rotor}}}{P_{\text{stator}}}$$

To improve the characteristics of the motor and its efficiency, a number of key issues need to be considered: selection of vibration modes and the consistency of the modal frequencies being used to generate the motion, the placement of the piezoelectric ceramic and the supporting plane, setting of the preload, and any influence from interfering modes [72]. One of the main issues associated with efficiency degradation is heat generation due to friction between the stator and rotor. Adding a preload, for instance, increases the efficiency of conversion from the stator to the rotor but it also increases the heat produced. A solution to this problem was proposed in [73] by selecting a proper lubricant. The efficiency becomes then dependent on the static preload and the lubricant type used. The maximum torque and efficiency show a linear relationship with the static preload [73]. Proper selection of the friction material can also improve the efficiency and the output power [15]. For instance, selection of PPS (Ryton) as a friction material can increase the power output and efficiency of the Shinsei USM motor (USR 30) to 7 W and 23 %, instead of 1 W and less than 5 %, respectively, obtained when PAMXD6 material is used instead. These developments clearly encourage other researchers to continue to investigate new friction materials for better output characteristics.

Table 2.6 Piezoelectric constant values for different materials [71]

Categories	PZT-4	PZT-5A	PZT-5H	PZT-8
Name	110	200-HD	610-HD	300-HD
d_{33} (pC/N)	360	400	750	235
d_{31} (pC/N)	−150	−190	−380	−100
d_{15} (pC/N)	520	550	920	390

The efficiency of USMs is defined as the ratio of the power transferred to the rotor to the stator power output [76]

Solutions to the temperature rise and heat transfer problems [14] are presented in [15] and [23]. The stator was mounted on a radiator to solve issues caused by temperature rise. In addition, decreasing the working temperature could be achieved by utilization of higher quality friction materials [23]. Another solution suggested by [15] is to use anti-resonant frequency instead of fundamental resonance frequency, which gives a higher mechanical quality factor and the same mechanical vibration level without generating heat. By applying this solution, conventional inexpensive power supply can be used, because of the need for high voltage and low current drivers, since the admittance is very low for anti-resonance state.

The future trend is still further miniaturization [23, 24, 71]. The main objective of many development works is to produce a compact, inexpensive USM, with higher power values and acceptable efficiency. Another challenge is to address the stator/motor contact problem and associated wear issues [14] and to enhance the understanding of the dynamics between the stator and rotor.

References

1. Y. Okazaki, Precision positioning control apparatus and precision positioning control method. U.S. Patent 5,801,939, 1998
2. W. Yao, M. Tomizuka, Robust controller design for a dual-stage positioning system, in *Proceedings of International Conference on Industrial Electronics, Control, and Instrumentation*, vol. 1, 1993, pp. 62–66
3. K. Tsai, J. Yen, Servo system design of a high-resolution piezo-driven fine stage for step-and repeat microlithography systems, in *Proceedings of Annual Conference of Industrial Electronics Society*, vol. 1, 1999, pp. 11–16
4. S. Kwon, W. Chang, Y. Youm, Robust and time-optimal control strategy for coarse/fine dual-stage manipulators, in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 4, 2000, pp. 4051–4056
5. S. Kwon, W. Chang, Y. Youm, On the coarse/fine dual-stage manipulators with robust perturbation compensator, in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 1, 2001, pp. 121–126
6. B. Zhang, Z. Zhu, Developing a linear piezomotor with nanometer resolution and high stiffness. IEEE/ASME Trans. Mechatron. **2**(1), 22–29
7. C. Zhao, *Ultrasonic Motors: Technologies and Applications* (Springer, Berlin, 2011)
8. B. Watson, J. Friend, L. Yeo, Piezoelectric ultrasonic micro/milli-scale actuators. Sens. Actuators A Phys. **152**(2), 219–233 (2009)
9. F. Ba-Tis, R. Ben-Mrad, A 3-DOF MEMS electrostatic piston-tube actuator. J. Microelectromech. Syst. **24**(4), 1173–1184 (2015)
10. D.K.C. Liu, J. Friend, L. Leo, A brief review of actuation at the micro-scale using electrostatics, electromagnetics and piezoelectric ultrasonics. Acoust. Sci. Technol. **31**(2), 115–123 (2010)
11. R. Ben-Mrad, H. Hu, Dynamic modeling of hysteresis in piezoceramics, in *Proceedings of IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, vol. 1, 2001, pp. 510–515
12. A. Henke, M.A. Kümmel, J. Wallaschek, A piezoelectrically driven wire feeding system for high performance wedge-wedge-bonding machines. Mechatronics **9**(7), 757–767 (1999)
13. J. Wallaschek, Ultrasonic motor research in Germany—past, present, future, in *Proceedings of the First International Workshop on Ultrasonic Motors and Actuators* (2005)
14. J. Wallaschek, Piezoelectric ultrasonic motors. J. Intell. Mater. Syst. Struct. **6**(1), 71–83 (1995)

15. K. Uchino, Piezoelectric ultrasonic motors: overview. Smart Mater. Struct. **7**(3), 273 (1998)
16. V.V. Lavrinenko, M. Nekrasov, Piezoelectric motor. Soviet Patent 217509 (1965)
17. H.V. Barth, Ultrasonic drive motor. IBM Technical Disclosure Bulletin **16**(7), 2263 (1973)
18. P.E. Vasiliev, V.S. Dvornin, A.V. Kondratiev, V.F. Kravchenko, U.S. Patent 4,240,141 (U.S. Patent and Trademark Office, Washington, DC, 1980)
19. T. Sashida, Trial construction and operation of an ultrasonic vibration driven motor. Oyo Butsiuri **6**(5), 713–718 (1982)
20. T. Sashida, Motor device utilizing ultrasonic oscillation. U.S. Patent 4,562,374 (U.S. Patent and Trademark Office, Washington, DC, 1985)
21. A. Kumada, A piezoelectric ultrasonic motor. Jpn. J. Appl. Phys. **24**(S2), 739 (1985)
22. Y. Ise, Ultrasonic motor. J. Acoust. Soc. Jpn. **54**, 6 (1987)
23. I. Prisacariu, C. C. Filipiuc, A general view on the classification and operating principle of piezoelectric ultrasonic motors. (2012 *International Conference and Exposition on Electrical and Power Engineering (EPE 2012)*, 25–27 October, Iasi, Romania)
24. K. Spanner, Survey of the various operating principles of ultrasonic piezomotors, in *Proceedings of the 10th International Conference on New Actuators*, June 2006
25. S.P. Salisbury, D.F. Waechter, R. Ben-Mrad, S.E. Prasad, R.G. Blacow, B. Yan, Closed-loop control of a complementary clamp piezoworm actuator. IEEE/ASME Trans. Mechatron. **12**(6), 590–598 (2007)
26. S. Salisbury, D.F. Waechter, R. Ben-Mrad, S.E. Prasad, R.G. Blacow, B. Yan, Complementary inchworm actuator for high-force, high-precision applications. IEEE/ASME Trans. Mechatron. **11**(3), 265–272 (2006)
27. P.E. Tenzer, R. Ben-Mrad, A systematic procedure for the design of piezoelectric inchworm precision positioners. IEEE/ASME Trans. Mechatron. **9**(2), 427–435 (2004)
28. D. Roberts, Development of a linear piezoelectric motor based on the inchworm model, in *1999 Symposium on Smart Structures and Materials* (International Society for Optics and Photonics, 1999), pp. 705–716.
29. J. Li, R. Sedaghati, J. Dargahi, D. Waechter, Design and development of a new piezoelectric linear Inchworm actuator. Mechatronics **15**(6), 651–681 (2005)
30. J. E. Frank, G. H. Koopmann, W. Chen, G. A. Lesieutre, Design and performance of a high-force piezoelectric inchworm motor, in *1999 Symposium on Smart Structures and Materials* (International Society for Optics and Photonics, 1999), pp. 717–723
31. T. Pandell, E. Garcia, Design of a piezoelectric caterpillar motor, in *Proceedings of ASME Aerospace Division*, vol. 52, 1996, pp. 627–648
32. S.P. Salisbury, R. Ben-Mrad, D.F. Waechter, S.E. Prasad, Design, modeling, and closed-loop control of a complementary clamp piezoworm stage. IEEE/ASME Trans. Mechatron. **14**(6), 724–732 (2009)
33. K. Uchino, *Piezoelectric Actuators and Ultrasonic Motors* (Kluwer Academic Publishers, Boston, 1997)
34. "L-104," Micro Pulse Systems Inc., [Online document], http://www.micropulsesystems.com, seen on 27 May 2002
35. http://www.physikinstrumente.com/technology/, seen on 5 March 2015
36. http://piezomotor.ports-it.net/technology/, seen on 5 March 2015
37. http://evolution.skf.com/, "Micromotor packs a punch", online article, seen on February 2015
38. J. Oliver, R. Neurogaonkar, J. Nelson, C. Bertolini, Rotary piezoelectric motor for vehicle applications. U.S. Patent 5,780,956 (1998)
39. K. Ohnishi, M. Umeda, M. Kurosawa, S. Ueha, Rotary Inchworm-type piezoelectric actuator. Electr. Eng. Jpn. **110**(3), 107–109 (1990)
40. K. Duong, E. Garcia,. Development of a rotary inchworm piezoelectric motor, in *Proceedings SPIE Smart Structures and Materials*, vol. 2445, 1995, pp. 782–788
41. S. Gursan, Development of a continuous-motion piezoelectric rotary actuator for mechatronics and micropositioning applications. MASc Dissertation, University of Victoria (1996)
42. K. Mori, Piezoelectric rotary actuator. U.S. Patent 4,468,583 (1984)

43. P.E. Tenzer, R. Ben-Mrad, On amplification in inchworm™ precision positioners. Mechatronics **14**(5), 515–531 (2004)
44. P. Tenzer, R. Ben-Mrad, Amplification in Inchworm precision positioners, in *1st joint Canada-US Workshop on Smart Materials and Structures* (St. Hubert, Quebec, Canada, 2001), pp. 77–84, 17–18 September 2001
45. S.P. Salisbury, R. Ben-Mrad, Analytical stiffness estimation for short flexures. Mechatronics **16**(7), 399–403 (2006)
46. T. Hemsel, J. Wallaschek, A piezoelectric linear vibration drive for high driving force. J Vibroengineering **1**, 7–12 (1999)
47. J. Zumeris, Ceramic motor. U.S. Patent 6,064,140 (2000)
48. O. Vyshnevsky, S. Kovalev, W. Wischnewskiy, A novel, single-mode piezoceramic plate actuator for ultrasonic linear motors. Ultrason. Ferroelectr. Freq. Control, IEEE Trans. **52**(11), 2047–2053 (2005)
49. T. Sashida, T. Kenjo, *An Introduction to Ultrasonic Motors* (Oxford Press, New York, 1993)
50. http://www.nanomotor.de/p_nanomotor.htm, seen on April 1, 2015
51. V. Klocke, Atomic precision and millimeter range, in *Feinwerktechnik, Mikrotechnik, Mikroelektronik*, vol. 104 (1996)
52. Y. Liu, W. Chen, J. Liu, S. Shi, A cylindrical standing wave ultrasonic motor using bending vibration transducer. Ultrasonics **51**(5), 527–531 (2011)
53. S. Park, S. He, Standing wave brass-PZT square tubular ultrasonic motor. Ultrasonics **52**(7), 880–889 (2012)
54. X. Lu, J. Hu, L. Yang, C. Zhao, A novel dual stator-ring rotary ultrasonic motor. Sens. Actuators A Phys. **189**, 504–511 (2013)
55. Y. Liu, J. Liu, W. Chen, S. Shi, A cylindrical traveling wave ultrasonic motor using longitudinal vibration transducers. Ferroelectrics **409**(1), 117–127 (2010)
56. Y. Liu, W. Chen, J. Liu, S. Shi, A cylindrical traveling wave ultrasonic motor using longitudinal and bending composite transducer. Sens. Actuators A Phys. **161**(1), 158–163 (2010)
57. Y. Liu, W. Chen, P. Feng, J. Liu, A rotary piezoelectric motor using bending vibrators. Sens. Actuators A Phys. **196**, 48–54 (2013)
58. Y. Liu, W. Chen, P. Feng, J. Liu, A square-type rotary ultrasonic motor with four driving feet. Sens. Actuators A Phys. **180**, 113–119 (2012)
59. X. Lu, J. Hu, L. Yang, C. Zhao, A novel in-plane mode rotary ultrasonic motor. Chin. J. Aeronaut. **27**(2), 420–424 (2014)
60. Y. Liu, W. Chen, J. Liu, S. Shi, A rotary ultrasonic motor using bending vibration transducers. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **57**(10), 2360–2364 (2010)
61. Y.X. Liu, J.K. Liu, W.S. Chen, X.H. Yang, A rotary ultrasonic motor using radial bending mode of ring with nested PZT excitation. J. Zhejiang Univ. Sci. A **13**(3), 189–196 (2012)
62. Y. Ting, J.M. Yang, C.C. Li, C.C. Yang, Y.C. Shao, P3P-6 modeling and design of a linear actuator by Langevin vibrators, in *Ultrasonics Symposium, 2006. IEEE* (2006), pp. 2337–2340
63. E. Moreno, P. Acevedo, M. Fuentes, M. Sotomayor, M. Borroto, M.E. Villafuerte, L. Leija, Design and construction of a bolt-clamped Langevin transducer, in *International Conference on Electrical and Electronics Engineering, Proceedings* (2005), pp. 393–395
64. T. Morita, M.K. Kurosawa, T. Higuchi, Cylindrical micro ultrasonic motor utilizing bulk lead zirconate titanate (PZT). Jpn. J. Appl. Phys. **38**(5S), 3347 (1999)
65. T. Morita, M. Kuribayashi Kurosawa, T. Higuchi, A cylindrical micro ultrasonic motor using PZT thin film deposited by single process hydrothermal method (/spl phi/2.4 mm, L= 10 mm stator transducer). IEEE Trans. Ultrason. Ferroelectr. Freq. Control **45**(5), 1178–1187 (1998)
66. J. Hu, K. Nakamura, S. Ueha, An analysis of a noncontact ultrasonic motor with an ultrasonically levitated rotor. Ultrasonics **35**(6), 459–467 (1997)
67. Y. Yamayoshi, S. Hirose, Improvement of Characteristics of Noncontact Ultrasonic Motor Using Acoustically Coupled Two Air Gaps. Jpn. J. Appl. Phys. **50**(7) (2011)
68. Y. Yamayoshi, J. Shiina, H. Tamura, S. Hirose, Noncontact ultrasonic motor with two flexural standing wave vibration disks. Jpn. J. Appl. Phys. **48**(9S1), 09KD10 (2009)

69. J. Lau, S.I. Gubarenko, R. Ben-Mrad, A novel plate-type linear ultrasonic motor with segmented electrodes, in *Proceedings of the 1st VMPT* (Montreal, QC, 2012)
70. X. Li, W.S. Chen, T. Xie, J.K. Liu, Novel high torque bearingless two-sided rotary ultrasonic motor. J. Zhejiang Univ. Sci. A **8**(5), 786–792 (2007)
71. C.Y. Lu, J.L. Li, W.Y. Pi, Ultrasonic motors using shear-type piezoelectric ceramics, in *2010 Symposium on Piezoelectricity, Acoustic Waves and Device Applications (SPAWDA),* IEEE (2010, December), pp. 465–469
72. Z. Li, C. Zhao, W. Huang, Z.L. Li, Several key issues in developing of cylinder type 3-DOF ultrasonic motor. Sens. Actuators A Phys. **136**(2), 704–709 (2007)
73. W. Qiu, Y. Mizuno, D. Koyama, K. Nakamura, Analysis of lubricating effect of hybrid *transducer*-type ultrasonic motor, in *Proceedings of 32nd Symposium on Ultrasonic Electronics* 32(2E4-3) (2011, November), pp. 301–302
74. S. He, P.R. Chiarot, S. Park, A single vibration mode tubular piezoelectric ultrasonic motor. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **58**(5), 1049–1061 (2011)
75. S.S. Jeong, T.G. Park, M.H. Kim, T.K. Song, Characteristics of a V-type ultrasonic rotary motor. Curr. Appl. Phys. **11**(3), S364–S367 (2011)
76. Y. Hojjat, M.R. Karafi, Introduction of roller interface ultrasonic motor (RIUSM). Sens. Actuators A Phys. **163**(1), 304–310 (2010)
77. T. Park, S. Jeong, H. Chong, K. Uchino, Design of thin cross type ultrasonic motor. J. Electroceramics **24**(4), 288–293 (2010)
78. G.L. Smith, R.Q. Rudy, R.G. Polcawich, D.L. DeVoe, Integrated thin-film piezoelectric traveling wave ultrasonic motors. Sens. Actuators A Phys. **188**, 305–311 (2012)
79. X. Lu, J. Hu, L. Yang, C. Zhao, Principle and experimental verification of novel dual driving face rotary ultrasonic motor. Chin. J. Mech. Eng. **26**(5), 1006–1012 (2013)
80. J.H. Hu, K. Nakamura, S. Ueha, Characteristics of a noncontact ultrasonic motor using acoustic levitation, in *Proceedings of Ultrasonics Symposium, 1996. IEEE*, vol. 1 (1996, November), pp. 373–376
81. Y. Yamayoshi, S. Hirose, Ultrasonic motor not using mechanical friction force. Int. J. Appl. Electromagn. Mater. **3**, 179–182 (1992)
82. S. Hirose, Y. Yamayoshi, H. Ono, A small noncontact ultrasonic motor, in *Proceedings of Ultrasonics Symposium, 1993. IEEE* (1993), pp. 453–456
83. J. Liu, B. Wu, Z. Yang et al., A new type of circular cylindrical non-contact ultrasonic motor. Acta Acustica **3**(2), 113–116 (2001) (In Chinese)
84. Y. Ji, C. Zhao, A new type non-contact ultrasonic motor with higher revolution speed. Piezoelectrics & Acoustooptics **28**, 527–533 (2006)
85. Y. Ji, C.S. Zhao, Cylinder type non-contact ultrasonic motor. J. Nanjing. Univ. Aeronaut. Astronaut. **37**(6), 690–693 (2005)
86. B. Yang, J. Liu, D. Chen, B. Cai, Theoretical and experimental research on a disk-type non-contact ultrasonic motor. Ultrasonics **44**(3), 238–243 (2006)

Chapter 3
Mechanical Design of High-Speed Nanopositioning Systems

Yuen Kuan Yong and Kam K. Leang

Abstract The performance of a nanopositioning system is tightly coupled to the quality of the mechanical design. Good mechanical design will minimize most position errors and improve overall accuracy and performance. Poor mechanical design, on the other hand, can lead to more errors than the issues associated with the electronics, control system, and other components. In this chapter, an overview of mechanical design is presented, where the emphasis on flexure-guided nanopositioning stages for high-speed nanopositioning. The discussions will focus on systems driven by piezoelectric actuators such as plate-stacks, which are readily available from a number of commercial suppliers. Design examples of parallel- and serial-kinematic scanners are presented to illustrate the design process.

3.1 Introduction

Nanopositioners are positioning devices that capable of providing motions with nanometer resolution. Nanopositioning systems have become the key enabling tools for precision motion applications in the fields of micro and nanotechnology [20, 104]. For example, the piezoelectric tube scanner (one type of nanopositioner) has helped to create one of the most versatile tools in nanotechnology—the atomic force microscope (AFM) [10, 11]. Since its invention, the AFM has brought about significant advances in many scientific and industrial fields including life sciences [6, 90], semiconductor manufacturing [91, 92], nano-machining, and nanofabrication [31, 55, 68, 85]. The AFM has revolutionized the way in which biologists analyze samples at resolution beyond the reach of optical microscopy [82].

Nanopositioners exist in many forms and a majority of them are driven by piezoelectric actuators. Examples include the aforementioned piezoelectric tube

Y.K. Yong (✉)
School of Electrical Engineering and Computer Science, The University of Newcastle, Callaghan, NSW 2308, Australia
e-mail: yuenkuan.yong@newcastle.edu.au

K.K. Leang
Department of Mechanical Engineering, The University of Utah, Salt Lake City, UT, USA
e-mail: kam.k.leang@utah.edu

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_3

scanner [57], flexure-based nanopositioners [35, 37, 69, 100, 104], piezoelectric benders [94], inchworm or stick-slip type nanopositioners [54, 62, 65], beam scanner [73], and mirror tilting positioner [83].

3.2 Scanning Probe Microscopy

A common application of nanopositioners is in the lateral and vertical positioning stages of scanning probe microscopes (SPMs) such as the AFM. Unlike a traditional optical microscope that uses light for imaging, an AFM image is formed by scanning a microcantilever probe over the surface, as illustrated in Fig. 3.1. The AFM is one of the most versatile microscopes due to its ability to work with conducting and nonconducting samples in a vacuum, air, or water [10]. The probe is a micromachined cantilever with a sharp tip protruding towards the sample surface. When the probe is brought into contact with the surface, the tip-to-sample interaction causes the cantilever to deflect vertically. This deflection is measured and used to construct an image of the sample. The AFM essentially "feels" the surface with a tiny, finger-like cantilever. In a vacuum, resolution of an AFM is on the order of 0.01 nm. With such high resolution, an AFM can generate topographical images of atoms, as well as to control, manipulate, and alter the properties of matter at the nanoscale.

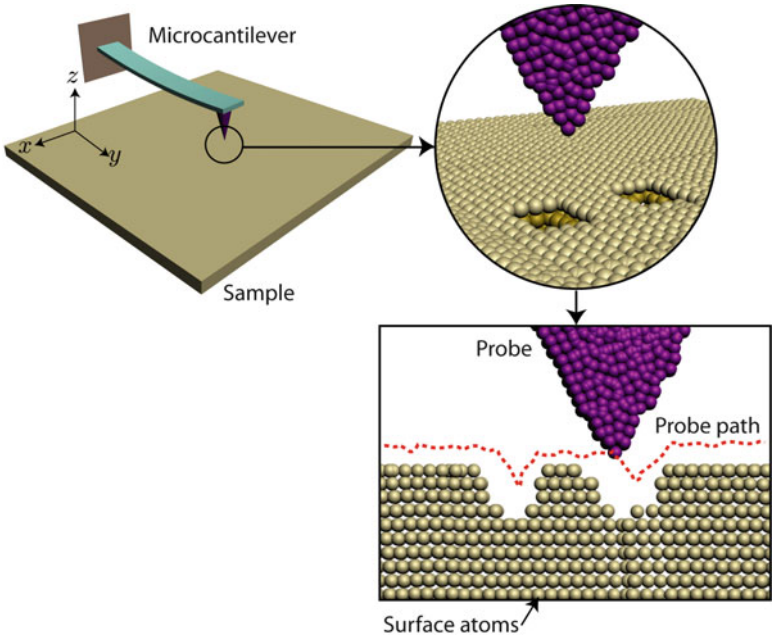


Fig. 3.1 The operation of an atomic force microscope

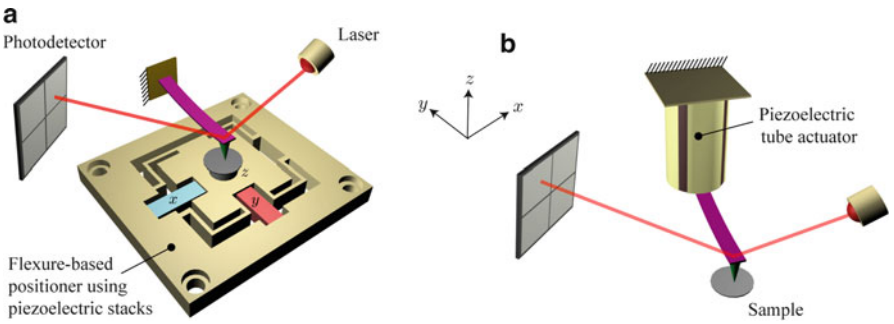


Fig. 3.2 Two positioning schemes for SPMs: (**a**) scan-by-sample and (**b**) scan-by-probe

The positioning of the probe tip relative to the sample can be achieved with two basic configurations: (a) scan-by-sample or (b) scan-by-probe as shown in Fig. 3.2. In the scan-by-sample configuration, the nanopositioner, flexure-based design shown equipped with three piezo stacks, moves the sample relative to a fixed probe. The x and y axes piezos position the sample along the lateral direction (parallel to the sample surface) and a z axis stack moves the sample vertically. The deflection of the cantilever is measured optically, by reflecting a laser beam off the end of the cantilever onto a nearby photodetector. Alternatively, in the scan-by-probe arrangement shown in Fig. 3.2b, a nanopositioner moves the probe relative to a fixed sample both laterally and vertically. In scan-by-probe systems, the laser and photodetector are required to move with the cantilever; however, this can be avoided by incorporating sensing elements into the cantilever itself, such as using piezoresistive, piezoelectric, or capacitive elements.

There are three basic operating modes of an AFM: contact, noncontact, and tapping mode. In contact mode, the probe interacts with the sample at very close range where the dominant force on the tip is repulsive. In this mode, the deflection of the AFM cantilever is sensed and a feedback controller is used to maintain a desired deflection. The spring constant of a contact mode AFM cantilever varies between 0.001 and 10 N/m.

Soft samples such as living cells have a contact stiffness comparable to, or less than, the cantilever stiffness, therefore, they may be deformed or damaged during contact mode operation. Noncontact mode avoids direct sample contact by exploiting attractive Van der Waals forces. In this mode, the AFM tip is hovered above the surface (at approximately 50–150 Å). By oscillating the tip, the effective stiffness of the microcantilever is effected by the force gradient of the attractive forces. The effective stiffness can be related to the sample topography by measuring or regulating the amplitude, phase, or resonance frequency of the probe. In general, noncontact mode AFM provides lower resolution than contact mode but does not pollute or damage the sample. Noncontact mode can also be used to measure long-range forces such as magnetic or electric fields in samples such as hard disk media or charged insulators.

For high-resolution imaging of soft samples such as living cells, polymers, and gels, tapping mode AFM is the preferred method. In this mode, the AFM cantilever is oscillated near its resonance frequency (50 kHz–1 MHz) using a piezoelectric actuator. As the AFM tip is brought into contact with the surface, the tip lightly touches or taps the surface. When the cantilever intermittently contacts the surface, the oscillating behavior is altered by the energy loss during the tip-to-sample interaction. The change in energy is monitored and used to construct an image of the surface.

Precision positioning is needed in many AFM applications. In particular, precise position control in both the lateral and vertical directions is needed to hold the probe at a desired location or to track a desired motion trajectory. For instance, when the AFM is used to create quantum dots (2–80 nm in size), accurate position control of the indenter tip is needed as the probe position directly affects the size, spacing, and distribution of the nanofeatures. Even 2–4 nm variation in size and spacing of the nanofeatures can drastically alter their properties [40]. Additionally, high-speed control of the probe's movement is needed for high throughput fabrication, imaging, and metrology. Without accurate motion control along a specific trajectory at high speed, oscillations can cause the tip to collide with nearby features, leading to excessive tip-to-sample forces and imaging artifacts. Large forces can damage the probe tip or soft specimens such as cells. Thus, accurate position control is critical in an AFM.

3.3 Flexure-Guided Nanopositioner

Flexure-guided nanopositioners have emerged as the design of choice for high-speed, high-accuracy nanoscale positioning. These nanopositioners have recently appeared in several commercial AFMs [72–74]. The key advantage of flexure-guided nanopositioners is their capability of high-speed scanning and low cross-coupling effects. With an increase in application for high-speed nanopositioning, such as for monitoring fast pace biological cell events [6] and high-speed nano-metrology [29], flexure-guided nanopositioners are becoming increasingly popular and critical, particularly for high-speed SPM systems [6, 37, 69, 104].

Flexure-guided systems exploit the advantages of compliant mechanisms (flexible mechanisms), where a flexible monolithic joint deforms elastically to offer accurate and repeatable fine motions. These flexible structures replace the traditional joints in rigid-link mechanisms, and offer motion guiding and mechanical amplification as illustrated in Fig. 3.3. Flexure-guided mechanisms are commonly machined from a single piece of material (monolithic) and no assembly of parts (i.e., between links and joints) is required. Consequently, the number of parts of a flexure-guided mechanism is substantially low, resulting in a low weight. These advantages enable the design of compact, light, and fast nanopositioners. Furthermore, the absence of moving and sliding joints provide a considerable benefit where problems such as wear, backlash, friction, and the need for lubrication are

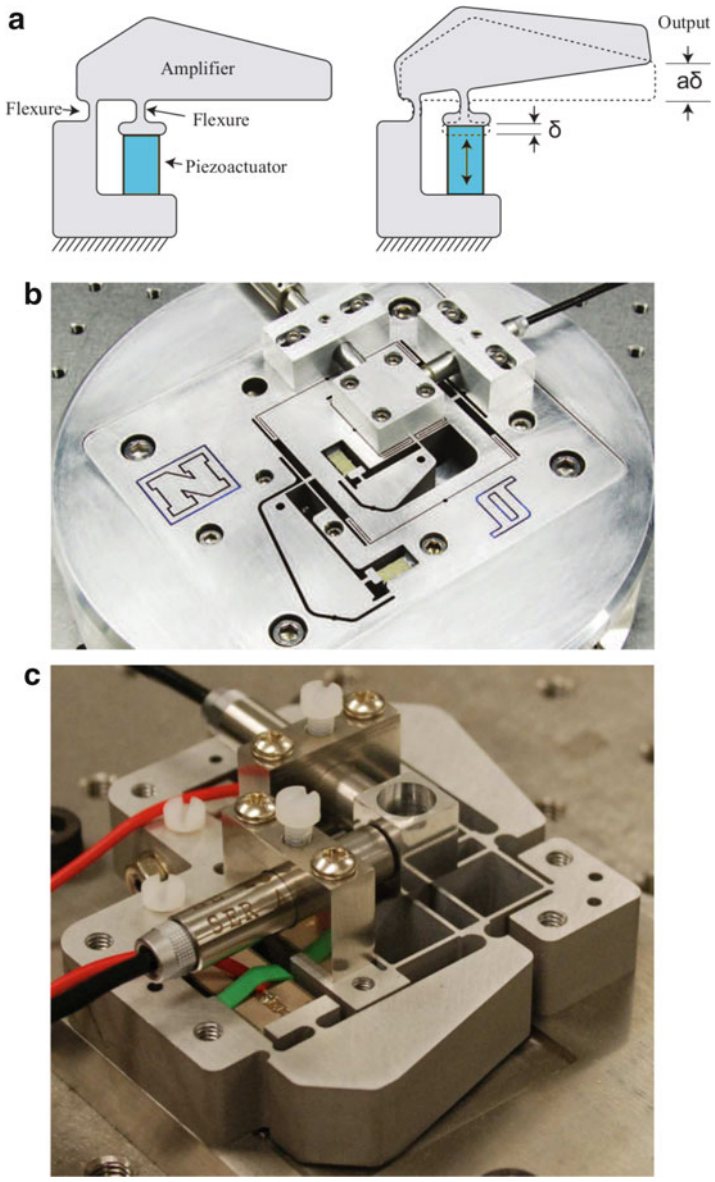


Fig. 3.3 Example compliant mechanism: (**a**) schematic of flexure and mechanical amplifier, (**b**) example flexure-based long-range serial-kinematic nanopositioner, and (**c**) example flexure-based long-range parallel-kinematic nanopositioner

eliminated [28, 44, 96]. Thus, flexure-guided nanopositioners provide repeatable, reliable, and smooth motions to fulfill the requirement of accurate nanoscale positioning for many nanoscale applications.

Flexure-guided mechanisms are categorized into two main configurations: serial or parallel (see examples in Fig. 3.4). Both serial- [3, 4, 35, 38] and parallel-kinematic [37, 49, 67, 69, 95, 96, 100] configurations have been used to design flexure-guided mechanisms. Depending on design requirements, each configuration has its own advantages and disadvantages. A serial-kinematic system is created (1) by stacking one piezo-stack actuator in series with another actuator to obtain the corresponding displacements [3], or (2) by nesting one flexure-guided nanopositioner into another [6, 35, 38]. Typically, high mechanical bandwidth can only be achieved in one motion axis [38]. This is mainly due to the relatively heavy weight that the base actuator would have to carry. Another disadvantage is the inability to measure and correct for parasitic motions caused by the cross-coupling effect from other axes [35]. Nevertheless, the cross-coupling effect can be minimized with carefully designed flexures and mechanisms [35, 38]. For raster scan purposes, the serial-kinematic design with one high-bandwidth stage is sufficient. Additionally, the serial-kinematic configuration is more cost effective as only one high-bandwidth, high-power piezo-amplifier is required [35].

Parallel-kinematic configurations have been used in commercially available nanopositioners [72, 74, 83]. Parallel structures offer high motion accuracy and high mechanical stiffness along the actuation direction, leading to high resonance frequencies. All actuators can be located relative to ground (base), thus reducing the inertia of the moving sample platform. Parallel structures generally have symmetrical configurations and they are less sensitive to temperature variations which could change the kinematics of the structures due to material expansion or contraction. All these characteristics of parallel structures make them attractive for accurate, fast nanopositioning applications. However, the cross-coupling between the x and y axes of parallel mechanisms is more difficult to deal with compared to serial mechanisms. Nevertheless, with properly designed flexures, nanopositioners with cross-coupling as low as -70 to -35 dB can be designed [41, 42, 83, 97, 100]. Although high-speed parallel-kinematic nanopositioners require high-power, high-bandwidth piezo-amplifiers for both x and y axes, the fast axis of a parallel nanopositioner can be chosen arbitrarily for rastering in AFMs [37, 69, 97, 100]. The parallel-kinematic configuration can also be used for emerging non-raster scan methods such as spiral-scan [52], cycloid-scan [102], and Lissajous-scan patterns [8, 88, 106].

3.4 Design Considerations

High-speed nanopositioners can be employed for video-rate SPM imaging, where the imaging rate can be as high as 30 frames per second [6, 38]. The frame rate establishes the required scan rate for the fast-axis motion. For example, for a 100×100 pixels image, the fast-axis scan rate should be at least 3 kHz. The parasitic motions along the x, y, and z axes should be within 1 % or less (≤ -40 dB). The resolution of the nanopositioner should be in the sub-nanometer

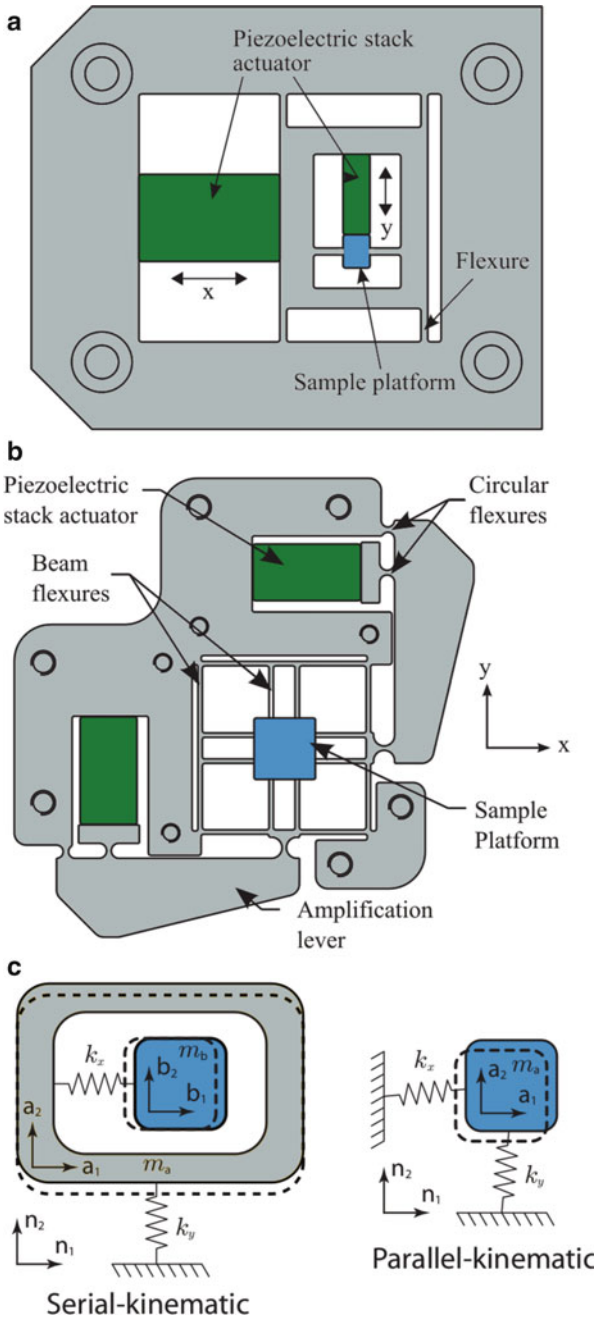


Fig. 3.4 Examples of serial- and parallel-kinematic nanopositioners. (**a**) A serial-kinematic system which is constructed by nesting one flexure-guided mechanism. The *y*-axis is the fast axis in this design. This is a reconstructed figure of Ando et al. [4]. (**b**) A parallel-kinematic system where all actuators are relative to ground. The fast axis can be either the *x*- or *y*-axis. The original figure can be found in [100]. (**c**) Lumped parameter models of both configurations

range. The scan range should be approximately $10 \times 10\,\mu m$ or more. In order to achieve the above characteristics and to avoid image artifacts, discussed above, it is necessary to consider (1) the mechanical stiffness of the nanopositioner and flexure design, (2) actuator properties and drive electronics, (3) material selection, and (4) manufacturing techniques.

3.4.1 Flexure Hinges

Flexures are important mechanical elements that commonly used in ultra-precision positioning devices to guide motions with nanometer accuracy. When forces are applied, these hinges generate motions by elastically deform their structures. Flexure hinges are increasingly popular for designs requiring monolithic manufacturing, smooth repeatable motions, low inertial mass, frictionless, and zero backlash.

3.4.1.1 Commonly Used Flexure Hinges in Nanopositioning

There are many types of flexures which provide unique properties to suit different positioning requirements. Examples of flexures which provide in-plane rotational degree-of-freedom (DOF) are basic beam flexure [69, 100], circular [44, 99, 104], corner-filleted [47], double-hinged [35, 89], and conical [47, 48, 80]. The out-of-plane stiffness of these flexures are high, however, if two degrees-of-freedom is desired, the depth of these flexures can be reduced to allow out-of-plane rotation (see Fig. 3.5b). There are also a multiple degrees-of-freedom flexure as illustrated in Fig. 3.5c where this hinge is free to rotate about all three axes [46].

Among all the aforementioned in-plane flexures, circular flexures provide the most accurate rotational motions because of the relatively small shift of the center of deflection. Circular flexures have a small in-plane rotational motion (bending), therefore, there are suitable for applications which require accurate positioning over a relatively small range [104]. For high displacement-orientated applications, corner-filleted flexures are more suitable due to their relatively low bending stiffness. However, there exhibit significant shift of center of deflection, making them less attractive for high precision-orientated applications [44, 104]. Flexures

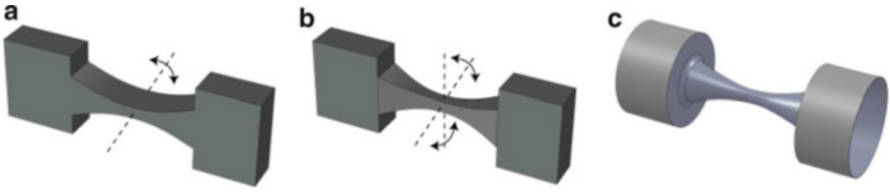


Fig. 3.5 Flexure degrees of freedom: (**a**) one, (**b**) two, and (**c**) multiple degrees-of-freedom

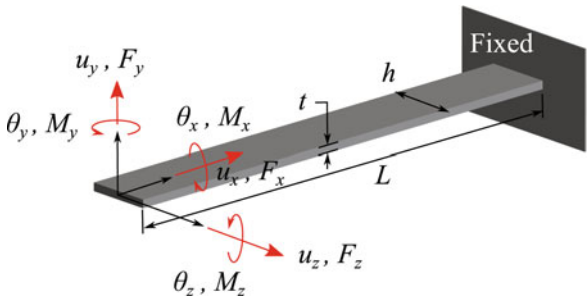


Fig. 3.6 A long, slender beam subjected to forces and moments

with a good combination of a large deflection angle and high rotational precision are the cross-axis pivot, cartwheel, and butterfly-type flexures [78, 93]. These types of flexures can be difficult to model due to the complexity of the structures.

Many methods can be used to derive a closed-form solution for the deformation of a flexure. One of these methods is the Castigliano's second theorem [44] (also known as Castigliano's displacement theorem). This method is a very effective method that allows calculation of the deformations of elastic bodies under the action of external and reactions loadings. The theorem is only valid for materials that deform elastically and obey Hooke's law for stress and strain.

According to the Castigliano's theorem, the linear displacement u_i and angular deformations θ_i of an elastic body at a point in a given direction ($i = x, y, z$) is the partial derivative of the total strain energy U with respect to the force F_i and M_i acting at that location, i.e.:

$$u_i = \frac{\partial U}{\partial F_i}, \tag{3.1}$$

$$\theta_i = \frac{\partial U}{\partial M_i}. \tag{3.2}$$

For a long, slender beam in Fig. 3.6 that is subjected to bending, shearing, axial load, and torsion, the total strain energy is [44]

$$U = U_{\text{bending}} + U_{\text{shearing}} + U_{\text{axial}} + U_{\text{torsion}}. \tag{3.3}$$

and

$$U_{\text{bending}} = U_{\text{bending},y} + U_{\text{bending},z} = \int_L \frac{M_y^2}{2EI_y} ds + \int_L \frac{M_z^2}{2EI_z} ds; \tag{3.4}$$

$$U_{\text{shearing}} = U_{\text{shearing},y} + U_{\text{shearing},z} = \int_L \frac{\alpha V_y^2}{2GA} ds + \int_L \frac{\alpha V_z^2}{2GA} ds; \tag{3.5}$$

$$U_{\text{axial}} = U_{\text{axial},x} = \int_L \frac{N_x^2}{2EA} ds; \tag{3.6}$$

$$U_{\text{torsion}} = U_{\text{torsion},x} = \int_L \frac{M_x^2}{2GJ} ds, \tag{3.7}$$

where M_i is the moment, V_i is the shear force, and N_x is the axial force. $G = \frac{E}{2(1+\nu)}$ is the shear modulus, I_i is the second moment of inertia, J is the torsional moment of inertia, α is the shear coefficient, and $\alpha = 6/5$ for a rectangular cross-section [16, 60, 107].

For a one-DOF, constant-width flexure hinge for which loadings are acting at location 1, as illustrated in Fig. 3.7, deformations at location 1 can be expressed into a matrix form by rearranging Eqs. (3.1)–(3.7) and factoring out all the common load parameters. This gives

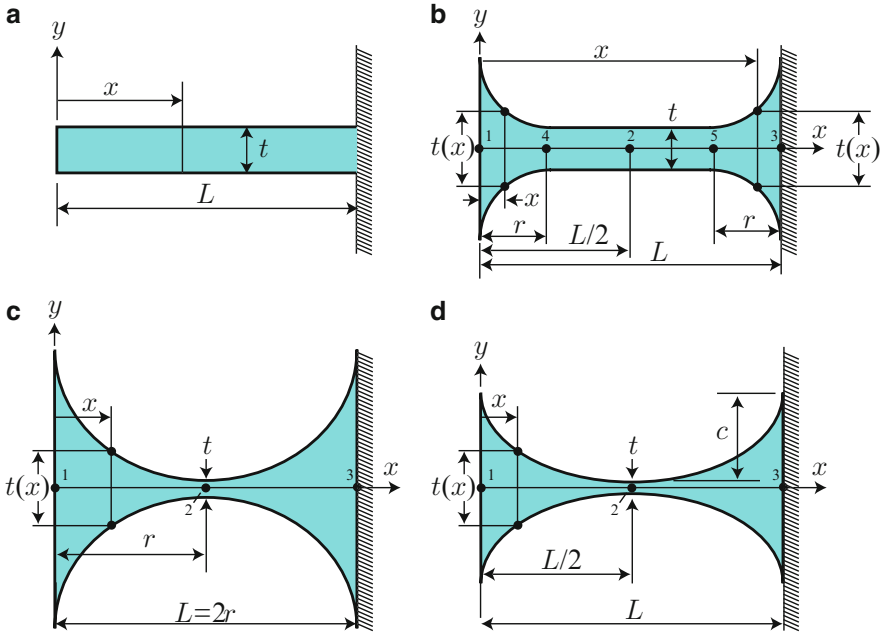


Fig. 3.7 Dimensions of commonly used flexures. (**a**) Basic beam, (**b**) corner-filleted, (**c**) circular, and (**d**) ellipse

$$\begin{Bmatrix} u_{1x} \\ u_{1y} \\ \theta_{1z} \\ u_{1z} \\ \theta_{1y} \end{Bmatrix} = \begin{bmatrix} C_{1,x-F_x} & 0 & 0 & 0 & 0 \\ 0 & C_{1,y-F_Y} & C_{1,y-M_z} & 0 & 0 \\ 0 & C_{1,\theta_z-F_y} & C_{1,\theta_z-M_z} & 0 & 0 \\ 0 & 0 & 0 & C_{1,z-F_z} & C_{1,z-M_y} \\ 0 & 0 & 0 & C_{1,\theta_y-F_z} & C_{1,\theta_y-M_y} \end{bmatrix} \begin{Bmatrix} F_{1x} \\ F_{1y} \\ M_{1z} \\ F_{1z} \\ M_{1y} \end{Bmatrix}, \tag{3.8}$$

where $C_{1,\theta_z-F_y} = C_{1,y-M_z}$ and $C_{1,\theta_y-F_z} = C_{1,z-M_y}$. The in-plane compliances are

$$C_{1,x-F_x} = \frac{1}{Eh}\int_0^L \frac{1}{t(x)}dx; \tag{3.9}$$

$$C_{1,y-F_y} = \frac{12}{Eh}\int_0^L \frac{x^2}{t(x)}dx + \frac{\alpha E}{G}C_{1,x-Fx}; \tag{3.10}$$

$$C_{1,y-M_z} = \frac{12}{Eh}\int_0^L \frac{x}{t(x)^3}dx; \tag{3.11}$$

$$C_{1,\theta_z-M_z} = \frac{12}{Eh}\int_0^L \frac{1}{t(x)^3}dx. \tag{3.12}$$

The out-of-plane compliances are

$$C_{1,z-F_z} = \frac{12}{Eh^3}\int_0^L \frac{x^2}{t(x)}dx + \frac{\alpha E}{G}C_{1,x-Fx}; \tag{3.13}$$

$$C_{1,z-M_y} = \frac{12}{Eh^3}\int_0^L \frac{x}{t(x)}dx; \tag{3.14}$$

$$C_{1,\theta_y-M_y} = \frac{12}{h^2}C_{1,x-F_x.} \tag{3.15}$$

Note that the torsional deformation θ_{1_x} is usually insignificant for nanopositioners discussed in this chapter. Therefore θ_{1_x} is neglected in Eq. (3.8). The in-plane compliance $C_{1,y-F_y}$ and out-of-plane compliance $C_{1,z-F_z}$ also include the shearing term $\frac{\alpha E}{G}C_{1,x-Fx}$. For a relatively long beam compared to its height, the shearing effect can be ignored according to the Euler–Bernoulli-beam model. This model assumes the planar cross section remains perpendicular to the neutral axis after the external bending has been applied. Shearing stresses and deformations are ignored. However, for a short beam, the shearing effect needs to be taken into consideration.

Deformation at location k can also be found by introducing a fictitious force F_k^* and moment M_k^* such that

$$u_k = \frac{\partial U}{\partial F_k^*}, \tag{3.16}$$

$$\theta_k = \frac{\partial U}{\partial M_k^*}. \tag{3.17}$$

However for the design analysis of most nanopositioners, deformations/stiffness at location 1 are the main design focus. Readers are referred to [44] for closed-form equations of deformations/stiffness at other locations of flexures.

Flexures that are commonly used in both serial- and parallel-kinematics nanopositioners are the planar types, i.e., basic beam, corner-filleted [44, 47], circular [44, 61, 87, 99], and elliptic [44, 80] as shown in Fig. 3.7. The closed-form compliance equations for the above flexure types can be found in [44]. These equations are presented here for convenience.

For a basic beam flexure with constant thickness, that is $t(x) = t$ for $0 \leq x \leq L$, as shown in Fig. 3.7a, the in-plane compliance equations are

$$\begin{aligned} C_{1,x-F_x} &= \frac{L}{Eht}; \\ C_{1,y-F_y} &= \frac{4L^3}{Eht^3} + \frac{\alpha L}{Ght}; \\ C_{1,y-M_z} &= \frac{6L^2}{Eht^3}; \\ C_{1,\theta_z-M_z} &= \frac{12L}{Eht^3}. \end{aligned} \tag{3.18}$$

The out-of-plane compliance equations are

$$\begin{aligned} C_{1,z-F_z} &= \frac{4L^3}{Eh^3t} + \frac{\alpha L}{Ght}; \\ C_{1,z-M_y} &= \frac{6L^2}{Eh^3t}; \\ C_{1,\theta_y-M_y} &= \frac{12L^3}{Eh^3t}. \end{aligned} \tag{3.19}$$

For a corner-filleted flexure as shown in Fig. 3.7b where the thickness is

$$t(x) = \begin{cases} t + 2\left[r - \sqrt{x(2r-x)}\right], & x \in [0, r] \\ t, \quad x \in [r, L-r] \\ t + 2\{r - \sqrt{(L-x)[2r-(L-x)]}\}, & x \in [L-r, r], \end{cases} \tag{3.20}$$

the in-plane compliances are

$$C_{1,x-F_x} = \frac{1}{Eh}\left[\frac{L-2r}{t} + \frac{2(2r+t)}{\sqrt{t(4r+t)}}\arctan\sqrt{1 + \frac{4r}{t} - \frac{\pi}{2}}\right];$$

$$C_{1,y-F_y} = \frac{3}{Eh}\left\{\frac{4(L-2r)(L^2 - Lr + r^2)}{3t^3}\right.$$

$$
\begin{aligned}
&+\frac{\sqrt{t(4r+t)}[-80r^4+24r^3t+8(3+2\pi)r^2t^2]}{4\sqrt{t^5(4r+t)^5}}\\
&+\frac{\sqrt{t(4r+t)}[4(1+2\pi)rt^3+\pi t^4]}{4\sqrt{t^5(4r+t)^5}}\\
&+\frac{(2r+t)^3(6r^2-4rt-t^2)\arctan\sqrt{1+\frac{4r}{t}}}{\sqrt{t^5(4r+t)^5}}\\
&+\frac{-40r^4+8Lr^2(2r-t)+12r^3t+4(3+2\pi)r^2t^2}{2t^2(4r+t)^2}\\
&+\frac{2(l+2\pi)rt^3+\frac{\pi t^4}{2}}{2t^2(4r+t)^2}+\frac{4L^2r(6r^2+4rt+t^2)}{t^2(2r+t)(4r+t)^2}\\
&-\frac{(2r+t)[-24(L-r)^2r^2-8r^3t+14r^2t^2+8rt^3+t^4]}{\sqrt{t^5(4r+t)^5}}\\
&\times\arctan\sqrt{1+\frac{4r}{t}}\Bigg\}\ ;
\end{aligned}
$$

$$
\begin{aligned}
C_{1,y-M_z}=&-\frac{6L}{Eht^3(2r+t)(4r+t)^2}\Bigg\{(4r+t)[L(2r_t)(4r+t)^2\\
&-4r^2(16r^2+13rt+3t^2)]+12r^2(2r+t)^2\sqrt{t(4r+t)}\\
&\times\arctan\sqrt{1+\frac{4r}{t}}\Bigg\}\ ;
\end{aligned}
$$

$$
\begin{aligned}
C_{1,\theta_z-M_z}=&\frac{12}{Eht^3}\Bigg\{L-2r+\frac{2r}{(2r+t)(4r+t)^3}\Bigg[t(4r+t)(6r^2+4rt+t^2)\\
&+6r(2r+t)^2\sqrt{t(4r+t)}\arctan\sqrt{1+\frac{4r}{t}}\Bigg]\Bigg\}\ .
\end{aligned}
\tag{3.21}
$$

The out-of-plane equations are

$$
\begin{aligned}
C_{1,z-F_z}=&\frac{12}{Eh^3}\Bigg\{\frac{(L-2r)(L^2-Lr+r^2)}{3t}+Lr\Bigg[\log\frac{t}{2r+t}-\frac{2(L-2r)}{\sqrt{t(4r+t)}}\\
&\times\arctan\sqrt{1+\frac{4r}{t}}\Bigg]\Bigg\}\ ;
\end{aligned}
$$

$$
C_{1,z-M_y}=\frac{6}{Eh^3t}\Bigg\{L(L-2r)+2r\Bigg[t\log\frac{t}{2r+t}-2(L-2r)\sqrt{\frac{t}{4r+t}}
$$

$$\times \arctan\sqrt{1+\frac{4r}{t}}\Bigg]\Bigg\};$$

$$C_{1,\theta_y-M_y} = \frac{6}{Eh^3t}\left[2L-4r=\pi t+4(2r+t)\sqrt{\frac{t}{4r+t}}\arctan\sqrt{\frac{t}{4r+t}}\right]. \quad (3.22)$$

For circular flexure hinge as shown in Fig. 3.7c, the thickness $t(x) = t + 2[r - \sqrt{x(2r-x)}]$. The in-plane compliances are

$$C_{1,x-F_x} = \frac{1}{Eh}\left[\frac{2(2r+t)}{\sqrt{t(4r+t)}}\arctan\sqrt{1+\frac{4r}{t}-\frac{\pi}{2}}\right];$$

$$C_{1,y-F_y} = \frac{3}{4Eh(2r+t)}\Bigg\{2(2+\pi)r+\pi t+\frac{8r^3(44r^2+28rt+5t^2)}{t^2(4r+t)}$$
$$+(2r+t)\sqrt{t(4r+t)}$$
$$\times\frac{-80r^4+24r^3t+8(3+2\pi)r^2t^2+4(1+2\pi)rt^3+\pi t^4}{\sqrt{t^5(4r+t)^5}}$$
$$-\frac{8(2r+t)^4(-6r^2+4rt+t^2)}{\sqrt{t^5(4r+t)^5}}\arctan\sqrt{1+\frac{4r}{t}}\Bigg\};$$

$$C_{1,y-M_z} = \frac{24r^2}{Eht^3(2r+t)(4r+t)^3}\Bigg[t(4r+t)(6r^2+4rt+t^2)$$
$$+6r(2r+t)^2\sqrt{t(4r+t)}\arctan\sqrt{1+\frac{4r}{t}}\Bigg];$$

$$C_{1,\theta_z-M_z} = \frac{C_{1,y-M_z}}{r}. \quad (3.23)$$

The out-of-plane compliance equations are

$$C_{1,z-F_z} = \frac{24r^2}{Eh^3}\log\frac{t}{2r+t};\quad C_{1,z-M_y} = \frac{C_{1,z-F_z}}{2r};$$

$$C_{1,\theta_y-M_y} = \frac{6}{Eh^3t}\left[4(2r+t)\sqrt{\frac{t}{4r+t}}\arctan\sqrt{\frac{t}{4r+t}}-\pi t\right]. \quad (3.24)$$

For an elliptical flexure as shown in Fig. 3.7d, with the thickness

$$t(x) = t + 2c\left[1-\sqrt{1-\left(1-\frac{2x}{c}\right)^2}\right], \quad (3.25)$$

where c is a constant, the in-plane compliances are

$$C_{1,x-F_x} = \frac{1}{4Ehc}\left[\frac{4(2c+t)}{\sqrt{t(4c+t)}}\arctan\sqrt{1+\frac{4c}{t}}-\pi\right];$$

$$C_{1,y-F_y} = \frac{3L^3}{16Eht^3c^3(2c+t)(4c+t)^2}\Bigg\{t[96c^5+96c^4t+8(11+4\pi)c^3t^2$$
$$+32(1+\pi)c^2t^3+2(2+5\pi)ct^4+\pi t^5]-4\sqrt{\frac{t}{4c+t}}(2c+t)^4$$
$$(-6c^2+4ct+t^2)\arctan\sqrt{1+\frac{4c}{t}}\Bigg\};$$

$$C_{1,y-M_z} = \frac{6L^2}{Eht^2(2c+t)(8c^2+t^2)}$$
$$\times\left[6c^2+4ct+t^2+\frac{6c(2c+t)^2}{\sqrt{t(4c+t)}}\arctan\sqrt{1+\frac{4c}{t}}\right]. \tag{3.26}$$

The out-of-plane compliances are

$$C_{1,z-F_z} = \frac{3L^3}{16Eh^3c^3}\Bigg\{2(4-\pi)c^2+4(1+\pi)ct+\frac{(2c+t)(-4c^2+4ct+t^2)}{\sqrt{t(4c+t)}}$$
$$\times\left[2\arctan\frac{2c}{\sqrt{t(4c+t)}}-\pi\right]\Bigg\};$$

$$C_{1,z-M_y} = \frac{3L^2}{2Eh^3c}\left\{\frac{2c+t}{\sqrt{t(4c+t)}}\left[2\arctan\frac{2c}{\sqrt{t(4c+t)}}+\pi\right]-\pi\right\}. \tag{3.27}$$

Double-hinged flexure as depicted in Fig. 3.8 has been widely used in nanopositioning systems to guide the displacement of the sample platform [32]. It is also used to restrain parasitic (out-of-plane and off-axis) motions so that the displacement stage only moves along the actuation axis. Recently a detailed design analysis of the double-hinged flexure for high-speed AFM serial-kinematic nanopositioning stage is reported in [35]. They pointed out that when a corner-filleted flexure is displaced in the vertical direction, the majority of the vertical displacement is caused by shear deformation of the center section of the flexure. Therefore, a simple way to increase the out-of-plane stiffness and consequently the resonance frequency is to thicken the center section of the beam flexure. This converts the beam flexure into a double-hinged flexure with a "rigid" center connecting link as shown in Fig. 3.8.

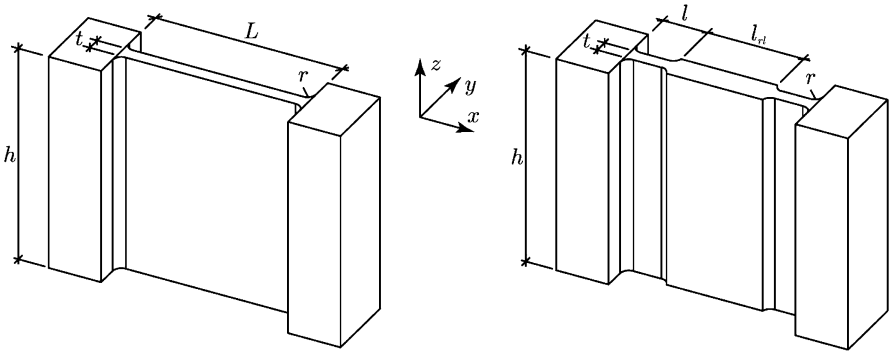


Fig. 3.8 Flexure design for increased out-of-plane stiffness. *Left*: conventional corner-filleted beam flexure; and *right*: serial-compliant double-hinged flexure with thickened center section

In this case, the cross-sectional area and second moment of inertia values are $A(x) = ht(x)$ and $I(x) = ht(x)^3/12$, respectively. For example, the thickness of the flexure in Fig. 3.8 is

$$t(x) = \begin{cases} t + 2\left[r - \sqrt{x(2r-x)}\right] & ,x \in [0,a] \\ t & ,x \in [a,b] \\ t + 2\left[r - \sqrt{(l-x)(2r-l+x)}\right] & ,x \in [b,c] \\ t + 2r & ,x \in [c,d] \\ t + 2\left[r - \sqrt{(l-g)(2r-l+g)}\right] & ,x \in [d,e] \\ t & ,x \in [e,f] \\ t + 2\left[r - \sqrt{g(2r-g)}\right] & ,x \in [f,L] \end{cases}, \tag{3.28}$$

where $a = r$, $b = l - r$, $c = l$, $d = L - l$, $e = d + r$, $f = L - r$, $g = L - x$, t and l are thickness and length of the thin section of the flexure, r is the fillet radius, $t + 2r = T$ is the thickness of the thickened section, and L is the length of the entire flexure. For this case, the compliance is determined by first determining the total strain energy while using the thickness function $t(x)$ in the area $A(x)$ and second moment of inertia $I(x)$ expressions. Again, the coordinate system is placed on the free end for simplification and to allow for direct integration. For instance, the total strain energy for bending due to a point load is

$$\begin{aligned} U &= \int_0^L \frac{M(x)^2}{2E\frac{ht(x)^3}{12}}\,dx + \int_0^L \frac{\alpha V(x)^2}{2Ght(x)}\,dx, \\ &= \frac{12F_y^2}{2Eh}\int_0^L \frac{x^2}{t(x)^3}\,dx + \frac{\alpha F_y^2}{2Gh}\int_0^L \frac{1}{t(x)}\,dx. \end{aligned} \tag{3.29}$$

Taking the partial derivative with respect to the applied force F_y gives the displacement

$$u_y = \frac{\partial U}{\partial F_y} = \frac{12F_y}{Eh}\int_0^L \frac{x^2}{t(x)^3}\,dx + \frac{\alpha F_y}{Gh}\int_0^L \frac{1}{t(x)}\,dx. \tag{3.30}$$

Then, the in-plane (and out-of-plane) stiffness is then calculated numerically by taking the ratio of the force to deflection.

3.4.1.2 Boundary Conditions of Flexures

The in- and out-of-plane compliance/stiffness is then calculated by taking the ratio of deflection to force/moment or vice versa in Eq. (3.8). Displacement and loading conditions of a flexure are dependent on its support conditions. The commonly used support conditions for beams are depicted in Fig. 3.9.

For example of a fixed-guided beam flexure (in which the moving end of the flexure remains parallel to the fixed-end) with concentrated end-load, which is a common boundary condition in nanopositioning design, the resultant moment is $-F_yL/2$. By using Eqs. (3.8), (3.18) and (3.19), the flexure displacement u_y due to applied force F_y and moment $M_z = -F_yL/2$ is

$$\begin{aligned} u_y &= C_{1,y-F_Y}F_y + C_{1,y-M_z}M_z \\ &= \left[\frac{4L^3}{Eht^3} + \frac{\alpha L}{Ght}\right]F_y + \left[\frac{6L^2}{Eht^3}\right]\frac{-F_yL}{2} \\ &= \left[\frac{L^3}{Eht^3} + \frac{\alpha L}{Ght}\right]F_y. \end{aligned} \tag{3.31}$$

Therefore, the stiffness k_y is

$$k_y = \frac{F_y}{u_y} = \left[\frac{L^3}{Eht^3} + \frac{\alpha L}{Ght}\right]^{-1}. \tag{3.32}$$

Similarly the vertical displacement u_z is

$$\begin{aligned} u_z &= C_{1,z-F_z}F_z + C_{1,z-M_y}M_y \\ &= \left[\frac{4L^3}{Eh^3t} + \frac{\alpha L}{Ght}\right]F_z + \left[\frac{6L^2}{Eh^3t}\right]\frac{-F_zL}{2} \\ &= \left[\frac{L^3}{Eh^3t} + \frac{\alpha L}{Ght}\right]F_z, \end{aligned} \tag{3.33}$$

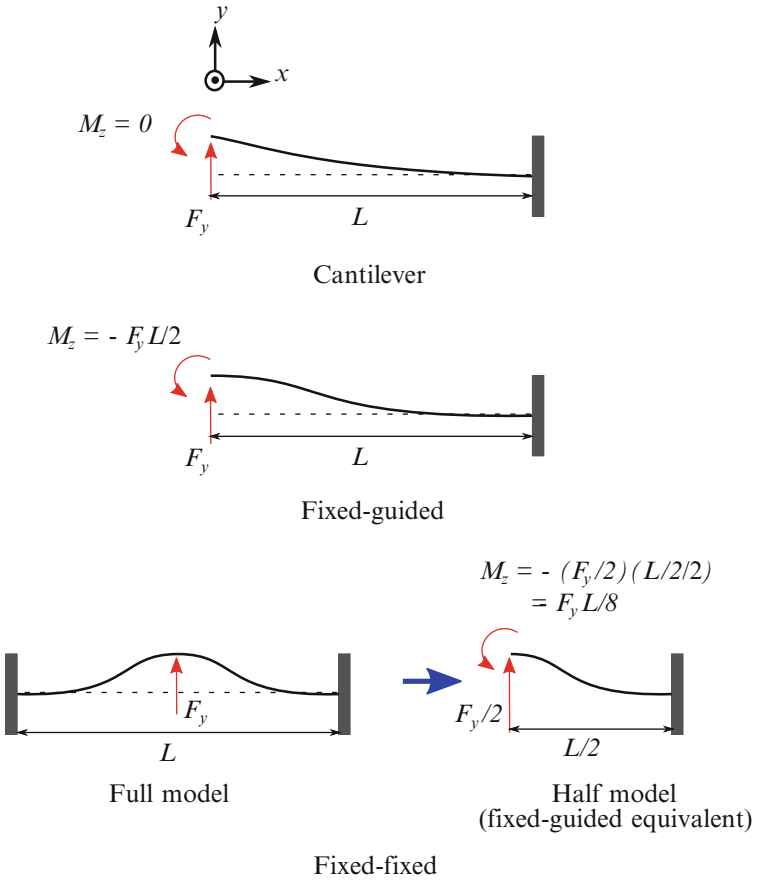


Fig. 3.9 Support conditions of a beam with concentrated end-load

and the stiffness k_z is

$$k_z = \frac{F_z}{u_z} = \left[\frac{L^3}{Eh^3t} + \frac{\alpha L}{Ght}\right]^{-1}. \tag{3.34}$$

The axial displacement u_x and stiffness k_x are

$$u_x = \frac{L}{Eht}F_x; \tag{3.35}$$

$$k_x = \frac{Eht}{L}. \tag{3.36}$$

Table 3.1 Stiffness equations of a basic beam flexure with loading and support conditions as shown Fig. 3.9

	Cantilever	Fixed-guided	Fixed-fixed
k_x	$\frac{Eht}{L}$	$\frac{Eht}{L}$	$\frac{4Eht}{L}$
k_y	$\left[\frac{4L^3}{Eht^3} + \frac{\alpha L}{Ght}\right]^{-1}$	$\left[\frac{L^3}{Eht^3} + \frac{\alpha L}{Ght}\right]^{-1}$	$\left[\frac{L^3}{16Eht^3} + \frac{\alpha L}{4Ght}\right]^{-1}$
k_z	$\left[\frac{4L^3}{Eh^3t} + \frac{\alpha L}{Ght}\right]^{-1}$	$\left[\frac{L^3}{Eh^3t} + \frac{\alpha L}{Ght}\right]^{-1}$	$\left[\frac{4L}{16Eh^3t} + \frac{\alpha L}{4Ght}\right]^{-1}$

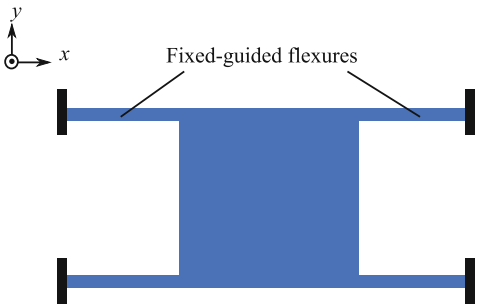


Fig. 3.10 A common design of a parallel-kinematic nanopositioner

Note that the ratio between the x- and y-axes stiffness (neglecting shearing) is $\frac{k_x}{k_y} = \left(\frac{L}{t}\right)^2$. This ratio is usually large for a long, slender flexure beam, which provides excellence reduction of orthogonal motion. The ratio between the z- and y-axes stiffness is $\frac{k_z}{k_y} = \left(\frac{h}{t}\right)^2$. Therefore, large flexure height h is one of the key factors for increasing the out-of-plane stiffness. However, it comes at an expense of a higher profile nanopositioner which increases the effective mass, thus reducing the actuation resonance. This effect may not be attractive for high-speed nanopositioning systems.

Using the similar method discussed above, stiffnesses of a basic beam for other support conditions, such as cantilever and fixed-fixed, are derived and exhibited in Table 3.1. The same procedures can be used to derive stiffnesses of flexures with other profiles as discussed in Sect. 3.4.1.

A common nanopositioning design constructed from connecting several flexures in parallel is shown in Fig. 3.10. The effective stiffness is Nk_i $(i = x, y, z)$, where N is the number of flexures. Generally speaking, the effective lateral stiffness k_x and k_y can be increased by (1) increasing the number of flexures and (2) decreasing the flexure length. The effective vertical stiffness k_z can be increased by (1) and (2), and also by using double-hinged flexures as previously discussed and shown in Fig. 3.8.

3.4.1.3 Commonly Used Flexures in MEMS

In linear micromachined structures such as gyroscopes [2] and MEMS nanopositioners [53, 56], the suspension systems are designed to be compliant (flexible) along the desired motion axes, and stiff along other directions. Most suspension systems utilize several fixed-guided flexures arranged in parallel as shown in Fig. 3.10 to

achieve the desired stiffness and resonance frequency. One disadvantage for the fixed-guided beam flexure is that, when experiencing large lateral deflection, the axial tension stress in the flexure increases significantly. This axial stress induces a nonlinear force-deformation relationship in the flexure [2]. To avoid this nonlinear stiffness behavior, a variation of the fixed-guided flexure, such as the folded (also known as U-shaped flexure), double-folded and crab-leg flexures are used to reduce the axial stress components [21, 63].

A folded flexure as shown in Fig. 3.11a is constructed by connecting two fixed-guided flexures in series. The two constrained points of the flexure are located at the same side, and the two unconstrained ends are free to expand or contract in all directions. Therefore, the lateral deflections does not result in axial loading. Since a folded flexure consists of two fixed-guided flexures in series, the stiffness of the folded beam is

$$\begin{aligned}\frac{1}{k_{y,\text{folded}}} &= \frac{1}{k_y} + \frac{1}{k_y} \\ &= \frac{2L^3}{Eht^3}, \\ k_{y,\text{folded}} &= \frac{Eht^3}{2L^3},\end{aligned} \tag{3.37}$$

where $k_y = Eht^3/L^3$ (neglecting shearing) as provided in Table 3.1. Similarly,

$$k_{x,\text{folded}} = \frac{Eht}{2L}; \tag{3.38}$$

$$k_{z,\text{folded}} = \frac{Eh^3t}{2L^3}. \tag{3.39}$$

For example shown in Fig. 3.11a, four folded flexures are arranged in parallel to create the suspension system. The effective stiffness of the system is $4\times k_{\text{folded}}$, i.e.:

$$K_{y,\text{folded}-\text{sys}} = \frac{2Eht^3}{L^3}; \tag{3.40}$$

$$K_{x,\text{folded}-\text{sys}} = \frac{2Eht}{L}; \tag{3.41}$$

$$K_{z,\text{folded}-\text{sys}} = \frac{2Eh^3t}{L^3}. \tag{3.42}$$

A double-folded flexure consists of two folded flexures connected symmetrically in parallel is shown in Fig. 3.11b. The stiffness of a double-folded beam is $2\times k_{z,\text{folded}}$, i.e.:

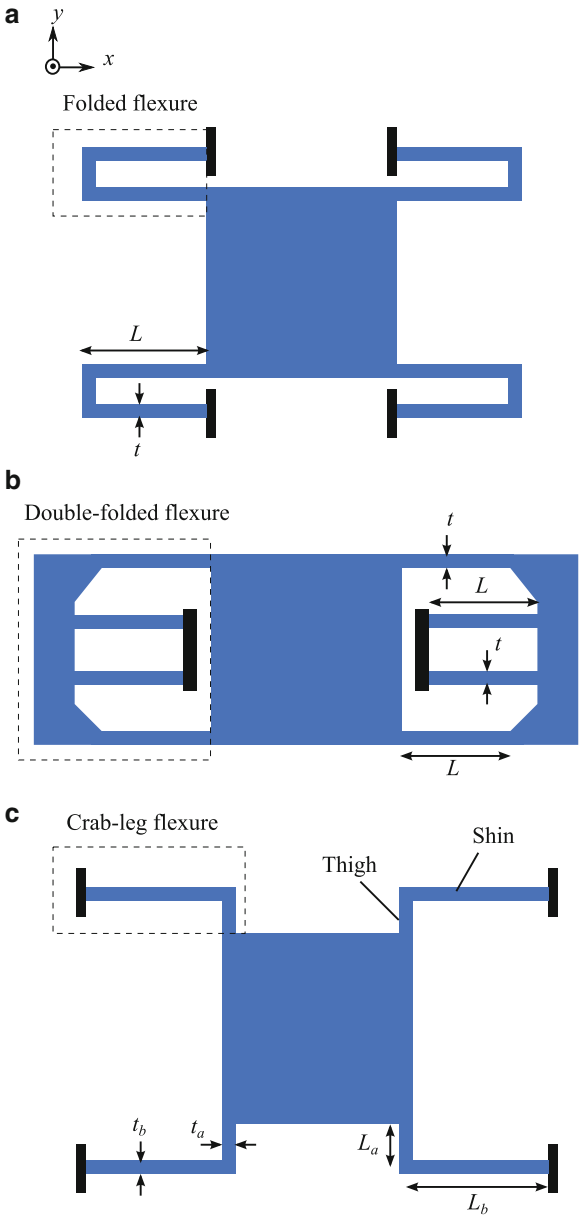


Fig. 3.11 MEMS suspension systems utilize (**a**) folded flexures, (**b**) double-folded flexures, and (**c**) crab-leg flexures

$$k_{y,\text{double-folded}} = \frac{Eht^3}{L^3}; \tag{3.43}$$

$$k_{x,\text{double-folded}} = \frac{Eht}{L}; \tag{3.44}$$

$$k_{z,\text{double-folded}} = \frac{Eh^3 t}{L^3}. \tag{3.45}$$

Note that for small deflection, the folded flexure provides half the stiffness of a fixed-guided flexure, while the double-folded flexure provides the same stiffness as that of the fixed-guided flexure.

For the suspension system that is constructed from two double-folded flexures as shown in Fig 3.11b, the effective stiffnesses are

$$K_{y,\text{df-sys}} = \frac{2Eht^3}{L^3}; \tag{3.46}$$

$$K_{x,\text{df-sys}} = \frac{2Eht}{L}; \tag{3.47}$$

$$K_{z,\text{df-sys}} = \frac{2Eh^3 t}{L^3}. \tag{3.48}$$

Another commonly used flexure in MEMS applications is a crab-leg flexure. The schematic of a crab-leg suspension system is shown in Fig. 3.11c. Crab-leg flexures are used in resonant-structure suspension systems such as MEMS gyroscopes [63]. The crab-leg flexure consists of a shin and a thigh segment. The thigh segment has length L_a and thickness t_a. The shin segment has length L_b and thickness t_b. The height of the flexure is h. For a suspension system constructed by four fixed-guided flexures (Fig. 3.10), both the k_x and k_y are scaled by the same term $4Eht/L$, therefore it is impossible to specify k_x and k_y independently. The crab-leg suspension system allows the two lateral stiffnesses k_x and k_y to take on values independent to each other [63]. The detailed stiffness derivation of the crab-leg suspension system can be found in [21]. Their stiffness equations are presented here for convenience.

$$K_{y,\text{crableg-sys}} = \frac{Eht_b^3 \left(4L_b + \gamma L_a\right)}{L_b^3 \left(L_b + \gamma L_a\right)}; \tag{3.49}$$

$$K_{x,\text{crableg-sys}} = \frac{Eht_a^3 \left(L_b + 4\gamma L_a\right)}{L_a^3 \left(L_b + \gamma L_a\right)}; \tag{3.50}$$

where $\gamma = \frac{I_b}{I_a} = \left(\frac{t_b}{t_a}\right)^3$.

3.4.1.4 Finite-Element-Analysis

Even though analytical expressions of flexures serve as a useful guideline in design, finite-element-analysis (FEA) is still a popular and powerful numerical tool for solving complicated engineering problems, including static and modal analyses, piezoelectricity, thermal analysis, coupled thermal-electric analysis, etc. Frequency response functions, which are useful for studying cross-coupling and out-of-plane

motions, can also be estimated using FEA package such as ANSYS. By careful consideration of mesh quality and convergence of results, accuracy within 15 % between FEA and experimental results can usually be achieved.

3.4.2 Piezoelectric Stack Actuators

3.4.2.1 Comparison of Piezo-Stack Actuated Nanopositioners

State-of-the-art flexure-guided piezo-driven stages [71], both direct drive serial-kinematic [5, 35, 38] and parallel-kinematic [70, 104, 105] configurations, have been developed with mechanical resonances in the tens of kHz range. These designs employ stiff and compact stack piezoactuators. The actuation resonance frequency for a stack piezoactuator held fixed on one end can be approximated by Kenton [33]

$$f_a = \frac{1}{2\pi L_p}\sqrt{\frac{3E}{\rho_{\text{piezo}}}}\ \text{Hz}, \tag{3.51}$$

where L_p is the length of the actuator, E is the elastic modulus, and ρ_{piezo} is the actuator's density. Substituting in values for $E = 33.9\,\text{GPa}$ and $\rho_{\text{piezo}} = 8000\,\text{kg/m}^3$ and assuming an achievable stroke (R) of 1 μm per mm of piezo length, it is shown that the maximum scanning frequency is $f_a = 567{,}460\,R^{-1}\,\text{Hz}\,\mu\text{m}$. When mass m_{end} is added to the free end of the actuator, the actuation resonance frequency reduces to

$$f_a = \frac{1}{2\pi}\sqrt{\frac{\frac{AE}{L_p}}{\frac{1}{3}AL_p\rho_{\text{piezo}} + m_{\text{end}}}}\ \text{Hz}, \tag{3.52}$$

where A is the cross-sectional area.

For comparison, the relationship between range and resonance frequency for a variety of commercial and custom nanopositioners is shown in Fig. 3.12 [33]. The range is plotted with respect to the resonance frequency for each stage when provided. When full details are not available for multi-axis positioners, it is assumed that the resonance frequency is provided for the stage with the largest displacement, and therefore; the largest range is plotted with respect to the lowest resonance frequency. Theoretically, the maximum attainable range R for a given actuation (longitudinal) resonance is $R = 567{,}460 f^{-1}$ (as derived in [33]). This value is for a fixed-free piezoactuator with a modulus of elasticity of 33.9 GPa and a density of $8000\,\text{kg/m}^3$ assuming 1 μm of travel per mm of piezo length. This relationship is shown in Fig. 3.12 as a dashed line. The commercial and custom nanopositioners in Fig. 3.12 are well below this theoretical limit. A trend line depicted as a solid line shows that for commercial and custom nanopositioners, the relationship is approximately $R = 30{,}688 f^{-0.916}$.

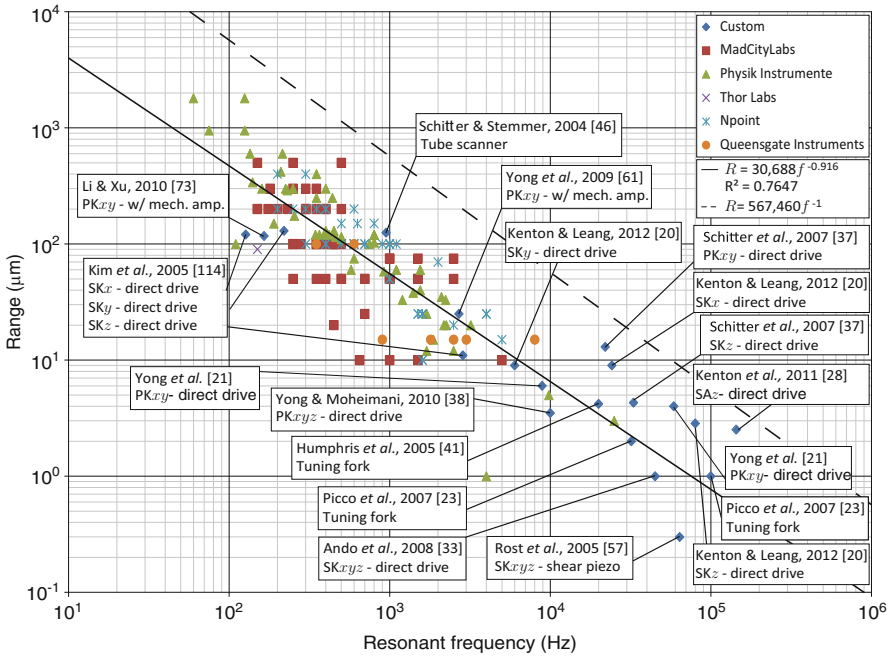


Fig. 3.12 High-performance commercial and custom nanopositioners plotted as range R with respect to resonance frequency f. The *solid line* represents a linear least-square-error line fit to the data points. The *dashed line* represents the theoretical first mechanical resonance in the actuation mode for a fixed-free piezoactuator (assuming 1 μm of travel per 1 mm length). *SK* serial-kinematic, *PK* parallel-kinematic, *SA* single axis, *x*, *y*, *z* refers to axis being referenced

Piezo-stack actuators are widely used to drive high bandwidth, flexure-guided nanopositioners, where repeatable sub-nanometer resolution is a requirement. Piezoelectric materials generally produce small strains, i.e., approximately 0.1 % of their original length. However, piezoelectric ceramic actuators can generate mechanical stresses in the order of tens of megapascals. That is, they can provide tens of newtons of force over a millimeter squared area. The response time of a piezoelectric actuator is determined by its molecular structure [39]. Due to the fact that molecular changes occur over a very small length scale, piezoelectric materials can respond very quickly to stimuli, e.g., an acceleration rate of 10^4 g can be achieved by a typical piezo-stack actuator [67]. Piezoelectric materials are quite suitable for applications that require fast actuation.

A piezo-stack actuator is constructed by "stacking" and gluing multiple thin layers of piezoelectric ceramics (as thin as 100 μm) together such that the polarization direction is aligned with the direction of stroke and the blocking force [83]. All layers are electrically connected in parallel. This enables the actuator to be operated at 100 V or less, and to achieve a stroke of approximately 0.1 % of the stack's length. The actuator also generates high actuation forces. The force

generated by a piezo-stack actuator is proportional to its d_{33} constant, as opposed to d_{31} that is smaller and often used in piezoelectric actuators. By combining a piezo-stack actuator with a flexure-guided amplification mechanism it is possible to realize nanopositioning platforms that can achieve a relatively large stroke with sub-nanometer positioning resolution [41, 43, 100].

3.4.2.2 Preloading

Because a piezo-stack actuator is made by gluing together multiple layers, the actuator is highly sensitive to tensile forces (pulling forces). Tensile forces can cause cracking of the ceramic material as well as separation of the layers. For a non-preloaded actuator, manufacturers often suggest that the tensile load should be within 5–10 % of the compressive load limit in order to avoid damaging the stack [83]. During high-speed operations, the piezo-stack experiences significant pushing and pulling forces due to the inertial forces cause by its own mass and the additional mass that it carries. A preload must be applied to compensate for the tensile force, thus protecting the piezo-stack actuator. Despite these challenges, stack actuators are used extensively in high-speed nanopositioning designs because of their high stiffness, fast response, and high force generation.

For flexure-guided mechanisms, a preload is commonly applied to the piezo-stack actuator using flexures (see Fig. 3.13). This is an elegant approach since it eliminates the need for additional components such as mechanical springs. The preload on the piezo-stack actuator is achieved when it is first installed into the flexure-guided structure by pressing the piezo-stack against the flexures using fasteners (see Fig. 3.13a) or metal shims. Other preloading methods include applying weights to deform the flexure elastically, allowing a piezo-stack actuator to be placed in its designated position. Weights are then released to restore the original position of the flexures (see Fig. 3.13b). The flexures will be slightly deformed and the resultant reaction force will hold the actuator in place.

The required preload can be calculated from the sample mass and its acceleration (computed from the known scan trajectory) using Newton's second law. A procedure for estimation of preload is discussed in [38].

3.4.2.3 Travel Range

The travel range of a piezo-stack actuator reduces when the stiffness of its flexures increases. One can imagine the flexures as a spring load (see Fig. 3.14) and the total available displacement of a piezo-stack actuator can be estimated as [83]

$$\Delta L = \Delta L_o \left(\frac{k_T}{k_T + k_s} \right), \tag{3.53}$$

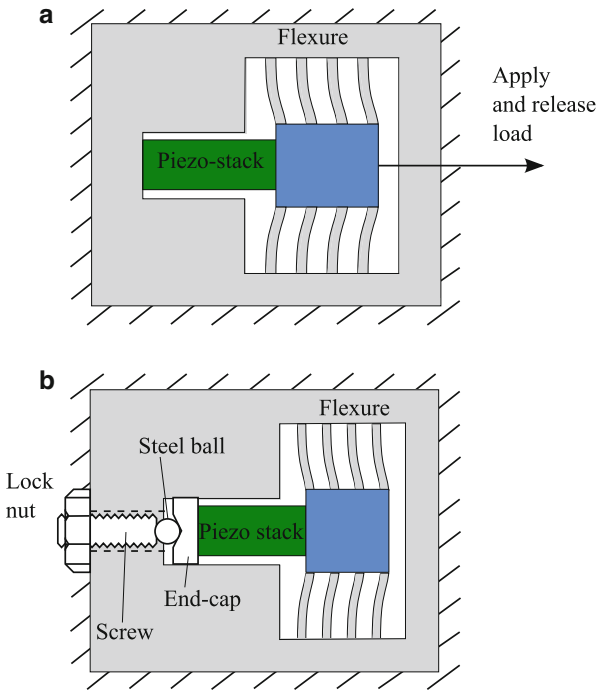
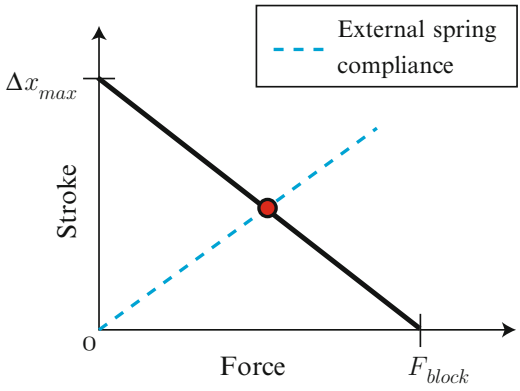


Fig. 3.13 Preloading a piezo-stack actuator. (**a**) Preloading using flexures. The deformation of the flexures are exaggerated in the diagram. (**b**) Preloading using fastener

Fig. 3.14 Displacement of a piezo-stack decreases as the spring load increases. Original diagram can be found in [83]



where ΔL is the displacement with external spring load, ΔL_o is the nominal displacement without external force, k_T is the piezo-stack's stiffness, and k_s is the external spring load. In most nanopositioner designs, this effect is insignificant since the stiffness of the flexures is designed to be 1/20th to 1/5th the stiffness of the piezo-stack actuator [38, 100, 104]. However, for nanopositioner designs where a large travel range is a priority, the consideration of total displacement in Eq. (3.53) may become significant.

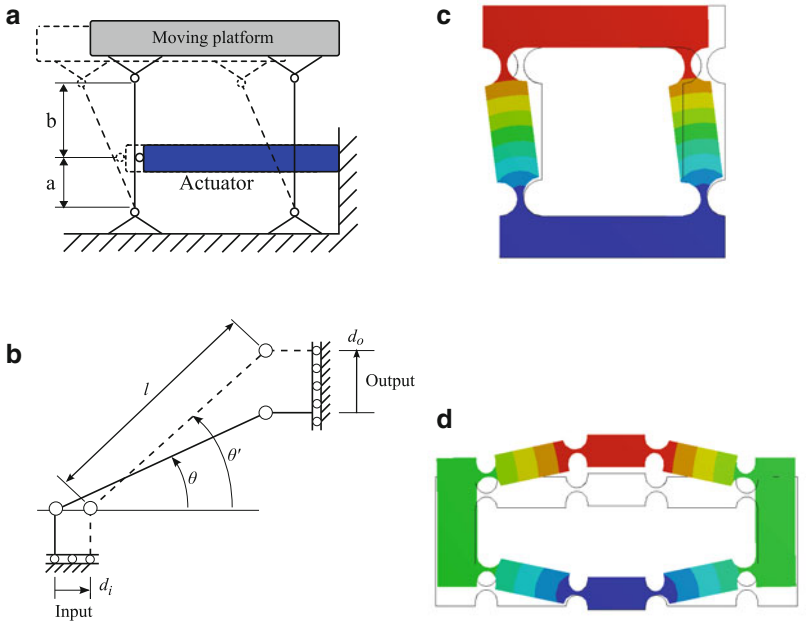


Fig. 3.15 Commonly used amplification mechanisms. (**a**) Schematic of a level-type mechanism. The amplification ratio can be estimated as $r = (a + b)/a$. (**b**) Schematic of a bridge-type mechanism. The amplification ratio can be estimated by $r = \left(\sqrt{l^2 \sin^2\theta + d_i\,(2l\cos\theta - d_i)} - l\sin\theta\right)/d_i$. (**c**) FE simulated displacement of the level-type mechanism. (**d**) FE simulated displacement of the bridge-type mechanism

Free stroke of most piezo-stack actuators is approximately 0.1 % of their length, which is insufficient for applications where a large displacement is a necessity. Multiple piezoactuators per degree-of-freedom have been used to increase travel range and bandwidth, but at the cost of increased power to drive the piezoactuators at high frequencies [5, 70]. Designs which involve mechanical amplification have been implemented to increase the range without having to increase the actuator's length [25, 26, 32, 36, 49, 71, 95, 100]. A parallelogram flexure guiding system as shown in Fig. 3.15a, c is a commonly used amplification mechanism. The amplification ratio of this mechanism can be estimated as $r = (a + b)/a$. The bridge-type amplification mechanism, illustrated in Fig. 3.15b, d, is also commonly used to amplify displacement of a piezo-stack actuator [36, 45, 50]. For this bridge-type amplification mechanism, the actuator is extended initially, and then retracted to provide the input displacement. The amplified output displacement of this mechanism is perpendicular to the direction of the actuator. The amplification ratio of this mechanism can be estimated as [45],

$$r = \left(\sqrt{l^2 \sin^2\theta + d_i\,(2l\cos\theta - d_i)} - l\sin\theta\right)/d_i. \tag{3.54}$$

Figure 3.15c, d illustrate the finite-element (ANSYS software) simulated displacements of the two flexure-guided amplification mechanisms, respectively.

Incorporation of a mechanical amplification mechanism into a nanopositioner is an effective way of increasing its achievable stroke. However, the added mass of the mechanical amplifier together with the flexible linkages lowers the mechanical resonance of the structure, which leads to a lower bandwidth. To achieve a high resonance frequency, the design of the nanopositioner has to be compact and rigid [35, 37]. Consequently, the length of the flexures and the amplification levers have to be small. However, short flexures and levers will reduce the overall motion of the device. This forces the designer to reach a compromise between the travel range and the bandwidth.

3.4.2.4 Self-Heating

During dynamic operation a piezoelectric actuator experiences self-heating, which is known to increase with actuation frequency and amplitude. It is also well known that electromechanical and electrical properties of a piezoelectric actuator can vary with temperature. Therefore, good thermal management is necessary for predictable performance, as well as to prevent premature failure and depoling of the piezoelectric material used. Heat management can be challenging due to the low thermal conductivity of the piezoelectric material. Creative heat sink designs can be employed to minimize the heating, but should not hinder the motion of the actuator. When an piezoelectric actuator is packaged in a metal case, the air gap can act as an insulator. Several vendors offer specialized actuator configurations, that are based on methods such as the "ThermoStable" technique, in order to improve heat management [7].

3.4.2.5 Electrical Considerations

Due to the highly capacitive nature of piezo-stack actuators, amplifiers with large current and power dissipation capabilities are needed to drive the actuators at high-speed [22]. Fleming [22] and Leang and Fleming [38] cover some of the electrical and drive issues related to high-speed operation of piezo-stack actuators. The current I_p and the associated power dissipation P_d in a linear amplifier can be estimated by Leang and Fleming [38]

$$I_p = sCV_p \tag{3.55}$$

$$P_d = I_p \left(V_s - V_p\right), \tag{3.56}$$

where s is the Laplace variable, C is the piezo-stack capacitance, V_p is the load voltage and, V_s is the supply voltage. For a nanopositioner that is to be operated at high speeds, selection of a piezo-stack actuator with a small capacitance is prudent. This will ensure that the actuator can be driven by an amplifier with sufficient

current and power dissipation capabilities. Typically, a smaller piezo-stack actuator possesses a smaller capacitance, but it has a lower stiffness and can provide a smaller stroke. These issues should be factored into the selection of a piezo-stack actuator for high-speed nanopositioning. Yong et al. [104] and Kenton [33] reports a comparison of piezo-stack actuators with various cross-sectional areas versus their maximum travel ranges and stiffness.

3.4.3 Material Considerations

The dynamic performance of a high-speed flexure-guided nanopositioner is relied on its material properties. Not only that, the choice of material for its supporting base also plays a major role in high-speed nanopositioner design. Material properties of great importance in dictating the dynamic performance of a nanopositioner are the Young's modulus of elasticity E, density ρ, the coefficient of thermal expansion (CTE). To facilitate the fabrication method of choice for ultra-precision machining in nanopositioning, that is the wire electrical-discharge-machining (WEDM), materials with sufficient electrical conductivity κ should also be taken into consideration during the design process. In summary, materials deemed suitable for realizing a high speed and thermally stable flexure-guided nanopositioner should have high E/ρ ratio, low CTE, and high κ.

Some commonly used materials in nanopositioning design are aluminum alloy Al 7075 and Al 6061, titanium alloy, invar alloy, and super invar alloy. Their properties are detailed in Table 3.2. Aluminum alloys such as Al 7075 and Al 6061 are relatively inexpensive and easily machined. They have been used to fabricate the vast majority of flexure-guided nanopositioners [17, 36, 49, 71, 100]. Aluminum has a high CTE, as stated in Table 3.2. Thus, it is not suitable for applications where time and temperature changes are the important factors. However, if the entire measurement can be completed in a short timeframe (i.e., in seconds or minutes) with adequate temperature control system, in a laboratory environment, aluminum alloys are deemed the most practical design material.

Table 3.2 A comparison of different material properties [15]

Material	E (GPa)	ρ (g/cm^3)	CTE (ppm/K)	E/ρ (GPa.cm^3/g)
Aluminum alloy (6061)	70	2.71	23.6	25.8
Aluminum alloy (7075)	72	2.81	23.6	26
Stainless steel	190	7.92	17.3	24
Titanium alloy (Grade 5)	106–114	4.42	8.8	24–26
Invar alloy (Invar 36)	148	8.1	1.3	18.4
Super Invar	148	8.1	0.3	18.4

Stainless steel is often used for fabricating the rigid base due to its high Young's modulus. Titanium alloys have a similar E/ρ value compared to aluminum alloys. They have a much lower CTE value. Thus, titanium alloys are suitable for applications where positioning accuracy needs to be maintained over a long period of time under varying temperature conditions. However, titanium is not always preferred over aluminum, because it is a more difficult material to machine [15].

Invar, a nickel–iron alloy, which exhibits even smaller values of CTE compared to titanium, can be used for applications with extreme thermal constraints. However, Invar alloys are considerably more expensive and somewhat difficult to machine [76]. These alloys are highly magnetic due to their large iron and nickel content [13]. Therefore, they are not suitable for applications where magnetic characteristic is a consideration.

3.4.3.1 Thermal Drift

Thermal expansion of the material from which a mechanism is made of can cause significant drift in its motion. Material expansion is inversely proportional to the melting point of a material [14]. The change of length (from l_0 to l_f) of a solid material for a change in temperature (from T_0 to T_f) is given by

$$\frac{l_f - l_0}{l_0} = \beta(T_f - T_0), \tag{3.57}$$

where β is the CTE and has units of $(°C)^{-1}$ or K^{-1}. For nanoscale motions, thermal effects cannot be ignored [58]. Careful material selection and good mechanical design are effective methods of minimizing thermal effects. For example, the CTE for aluminum is $23 \times 10^{-6}/°C$, while for Super Invar alloy is only $0.3 \times 10^{-6}/°C$, over 70-times lower. Effective practices also include carefully matching the stage material with the materials of surround support structures. Also, materials with high thermal conductivity quickly reach thermal equilibrium, thus minimizing transient distortion during thermal expansion.

3.4.4 Manufacturing Techniques

The appropriate manufacturing technique for a given design depends on the scale of the positioning stage and the material from which the mechanism is to be machined. The positioning accuracy of a nanopositioner depends on the precision with which its flexures are machined [66]. In particular, it is crucial to ensure all cuts of flexures have parallel sides without tapering. For example, the flexures of the nanopositioner in Fig. 3.16 have to be as identical as possible to avoid parasitic motions (or out-of-plane motions), caused by uneven stiffness of flexures.

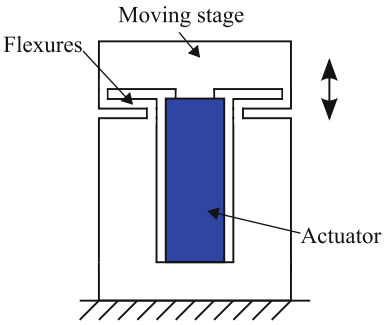


Fig. 3.16 Single axis nanopositioner. Flexures located at both sides of the actuator have to be as identical as possible to avoid parasitic motions due to asymmetric flexural stiffness

Standard milling, turning, and drilling techniques are best suited for metals, such as aluminum, titanium, and steel. These techniques are best for feature sizes above 1 mm, and they can achieve tolerances on the order of ± 0.001 in ($\approx \pm 25\,\mu$m). Some of the high-precision machining methods that are often employed to construct nanopositioning platforms are discussed here.

3.4.4.1 Wire Electrical-Discharge-Machining

Wire Electrical-Discharge-Machining (EDM) is one of the most frequently used machining techniques for fabricating monolithic structures such as flexure-guided nanopositioners, owing to its ability to machine complicated shapes with relatively fast cutting speed and accuracy [27, 64]. This method of machining was developed in the 1940s and is based on the erosion of a metallic material in the path of electrical discharges that form an arc between a continuously moving electrode tool (wire) and the workpiece. Dielectric fluid (deionized water) is ejected into the sparking area to flush away the eroded particles [64, 81]. Wire diameters of approximately 100 μm are often used. Traditional machining techniques are used to remove the bulk of the stock before performing EDM. Dimensional accuracy on the order of $\pm 12\,\mu$m can be achieved using the EDM process without severely tapering sides and a good surface finish (1.95 Ra or less) [1]. Note that Ra is the arithmetic mean of the absolute roughness profile ordinates [79].

3.4.4.2 Diamond Tool Machining

This is an ultra high-precision machining method which provides high geometrical accuracy, good surface quality, capability to machine complicated profiles and can be used to cut very hard materials [75]. However, this machining process can be labor-intensive, leading to high fabrication cost.

3.4.4.3 Laser Machining

Laser machining operations have large production rates. Therefore, they are used extensively in mass-production of industrial apparatus. The laser machining technique operates on the simple principle of heating and melting of workpieces. However, there are possible defects which could affect the performance of a nanopositioner. For example, cuts may not have parallel sides, a workpiece may have recast layers, i.e., materials on both sides of the cut may re-solidify after cooling, leading to different properties for these layers, and cracks may appear on surface of a workpiece due to poor surface finishing [19].

3.5 Design Example of High-Speed Lateral Nanopositioners

3.5.1 Parallel-Kinematic Nanopositioner

This section presents a design example of a high-speed parallel-kinematic nanopositioner as depicted in Fig. 3.17. Due to its mechanical symmetry in *xy*-plane, the nanopositioner can be operated in both raster and non-raster scan modes [51, 102, 106]. A piezoelectric stack actuator is used to drive each motion axis of the nanopositioner. Left-spring or beam-like flexures are used to guide the sample stage along the direction of motion of the actuator, while restraining other parasitic motions.

To preload the *x* and *y* piezoelectric stacks, weights are first applied to deform the flexure elastically, allowing a piezoelectric stack actuator to be placed in its designated position. Weights are then released to restore the original position of the flexures. The procedure results in the application of the necessary preload to the piezoelectric stack.

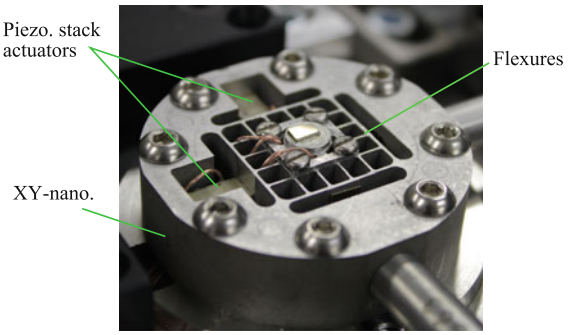


Fig. 3.17 A parallel-kinematic *xy*-nanopositioning stage

3.5.1.1 Effective Stiffness and Displacement Estimations

A key design specification for a high-speed AFM nanopositioner is a very high first mechanical resonance frequency along the actuation direction [35, 100]. This requirement, however, often leads to a reduction in the travel range of the stage. This is due to the inevitable increase in the structural stiffness when the resonance frequency is increased. This trade-off is best described by $\omega_n = \sqrt{k/m}$, where k and m represent the effective stiffness and mass of the stage, respectively, and ω_n is the resonance frequency of the mechanical system. The following study of the effect of flexure dimensions on the nanopositioner's resonance frequency and displacement was instrumental in reaching a compromise between the speed and travel range of the stage.

The stiffness of individual flexure was derived using strain energy method [9] and Castigliano's second theorem [44, 47]. These methods have been commonly used in flexure design literature [35, 46, 95, 96]. The in-plane displacement of a beam flexure can be defined as,

$$\begin{bmatrix} u_x \\ u_y \\ \theta_z \end{bmatrix} = \begin{bmatrix} C_{x,F_x} & 0 & 0 \\ 0 & C_{y,F_y} & C_{y,M_z} \\ 0 & C_{\theta_z,F_y} & C_{\theta_z,M_z} \end{bmatrix} \begin{bmatrix} F_x \\ F_y \\ M_z \end{bmatrix}, \tag{3.58}$$

where the displacements and loads are as shown in Fig. 3.6. The compliance values are derived in [35], as $C_{x,F_x} = L/AE$, $C_{y,F_y} = L/3EI + \alpha L/GA$, $C_{y,M_z} = C_{\theta_z,F_y} = L^2/2EI$, $C_{\theta_z,M_z} = L/EI$. The corresponding notation is described in Table 3.3.

For a fixed-guided beam, the resultant moment M_z at the guided end of the beam due to the applied F_y at that point is $M_z = -F_y L/2$. Note that the proposed flexures used in this work are short flexures. Therefore, the shear-induced deformation term $\alpha L/GA$ in C_{y,F_y} should not be ignored. Inclusion of the shear term is known to result in different deformations in flexures [99]. From Eq. (3.58), the displacement of the flexure along the y-direction is obtained as:

Table 3.3 Notation

Symbol	Definition
t	Beam thickness
L	Beam length
h	Beam height
A	Cross-sectional area, $A = ht$
E	Young's modulus
v	Poisson's ratio
G	Shear modulus, $G = E/[2(1+v)]$
I	Second moment of inertia about the z-axis, $I = ht^3/12$
J	Polar moment of inertia, $J = ht\left(h^2 + t^2\right)/12$
α	Shear correction factor

$$u_y = C_{y,F_y} F_y + C_{y,M_z} M_z$$
$$= F_y \left(\frac{L^3}{3EI} + \frac{\alpha L}{GA} \right) - \frac{F_y}{2} \left(\frac{L^3}{2EI} \right). \tag{3.59}$$

Substituting $I = ht^3/12$ and $A = ht$ into Eq. (3.59), and rearranging the equation, the stiffness of the flexure along the y-direction is determined as:

$$k_y = \frac{F_y}{u_y} = \left[\frac{1}{Eh} \left(\frac{L}{t} \right)^3 + \frac{\alpha}{Gh} \left(\frac{L}{t} \right) \right]^{-1}. \tag{3.60}$$

Since we know E, G, and h, the parameter that determines the flexure stiffness is L/t. Flexures of the xy-nanopositioner in Fig. 3.17 are arranged in parallel. Therefore, the general stiffness equation of the xy-nanopositioner along their direction of actuation is $k_s = N k_y$, where N is the number of flexures. The maximum displacement of the nanopositioner D_s is dependent on the piezoelectric stack actuator and it is estimated as:

$$D_s = \Delta l \left(\frac{k_p}{k_p + k_s} \right), \tag{3.61}$$

where Δl is the maximum stroke of the unloaded piezoelectric stack actuator and k_p is the stiffness of the piezoelectric actuator.

3.5.1.2 Resonance Frequency Estimation

The lateral resonance frequencies along the direction of actuation are estimated using the algebraic eigenvalue method [30]. Assuming the system is a two-degree-of-freedom undamped system as shown in Fig. 3.18, we may write

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{0}, \tag{3.62}$$

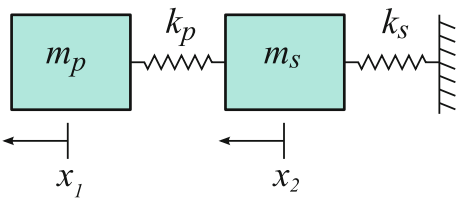


Fig. 3.18 Each actuation axis of the xy-nanopositioner is modeled as a two-degree-of-freedom undamped system. m_p and k_p are the mass and stiffness of the piezoelectric stack actuator. m_s and k_s are the mass and stiffness of the nanopositioning nanopositioner

where $\mathbf{M} = \begin{bmatrix} m_p & 0 \\ 0 & m_s \end{bmatrix}$, and $\mathbf{K} = \begin{bmatrix} k_p + k_s & -k_s \\ -k_s & k_s \end{bmatrix}$. m_p and k_p are the mass and stiffness of the piezoelectric stack actuator. Also, m_s and k_s are the mass and stiffness of the nanopositioner. Substituting $\mathbf{x} = \mathbf{M}^{-1/2}\mathbf{q}$ into Eq. (3.62) and multiplying the resulting equation by $\mathbf{M}^{-1/2}$ yields

$$\mathbf{I}\ddot{\mathbf{q}}(t) + \mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2}\mathbf{q}(t) = \mathbf{0}. \tag{3.63}$$

Let $\tilde{\mathbf{K}} = \mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2}$, by substituting $\mathbf{q}(t) = \mathbf{v}e^{j\omega t}$, where $\mathbf{v}$ is a constant vector, into Eq. (3.63) we obtain

$$\tilde{\mathbf{K}}\mathbf{v} = \omega^2\mathbf{v}. \tag{3.64}$$

If we replace ω^2 with λ in Eq. (3.64), the solutions for λ are the eigenvalues of the matrix $\tilde{\mathbf{K}}$. The resonance frequencies are $f_1 = \sqrt{\lambda_1}/2\pi$ and $f_2 = \sqrt{\lambda_2}/2\pi$. Note that the smaller of the two frequencies f_1 and f_2 will be the first translational mode of the system and is denoted as f_{res}.

3.5.1.3 Parametric Study

In this section, a parametric study of the *xy*-nanopositioner design is presented. In order to have a quantitative analysis of the *xy*-nanopositioner design, the relationship among N, L/t, k_s, D_s, and f_{res} is plotted in Fig. 3.19. The number of flexures, N is stepped from 1 to 20 and L/t is varied from 3 to 15.

As the ratio of L/t increases, the stiffness and resonance frequency decrease while the displacement increases. This is an unavoidable trade-off between stiffness/resonance frequency and displacement. Intuitively, the resonance frequency increases when the number of flexures N increases. However, the additional mass of each flexure lowers the resonance frequency. For example, as shown in Fig. 3.19, at $L/t = 6$ and $N = 10$, f_{res} is increased to 12 kHz, which is a threefold increment in resonance frequency compared to that of a single flexure ($N = 1$). However, when N is increased to 20, f_{res} is improved only by a factor of four at the expense of the displacement range being reduced by 20 %. Therefore, the resonance frequency of the nanopositioner may not be increased significantly by introducing more flexures into the design.

For the *xy*-nanopositioner design, K_s was set to approximately 10 N/μm to ensure that adequate preload was applied to the piezoelectric stack actuator. To achieve approximately 10 kHz resonance frequency without sacrificing too much displacement, $L/t = 9$ and $N = 12$ were selected in the design.

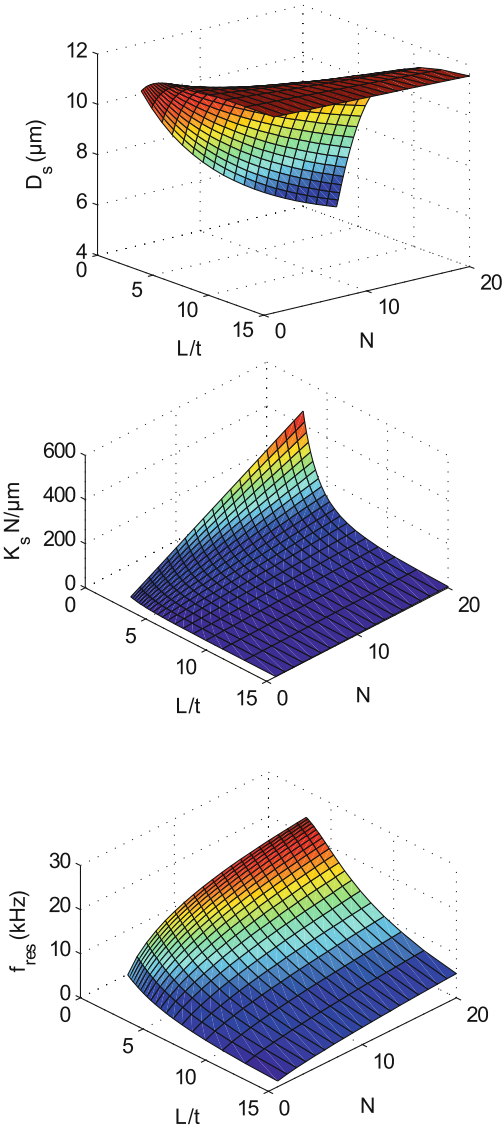


Fig. 3.19 Displacement D_s, stiffness K_s, and resonance frequency f_{res} plots of the *xy*-nanopositioner in relation to the ratio L/t and the number of flexure N

3.5.1.4 Finite-Element-Analysis

The analytical equations derived in Sects. 3.5.1.1 and 3.5.1.2 were based on a single axis mass-spring model. They serve as a useful tool for initial parametric design study of the nanopositioner. In this section, a finite-element (FE) model of the *xy*-nanopositioner is presented. The FE simulations were used to obtain a better estimation of the stiffness and resonance frequencies of the nanopositioner.

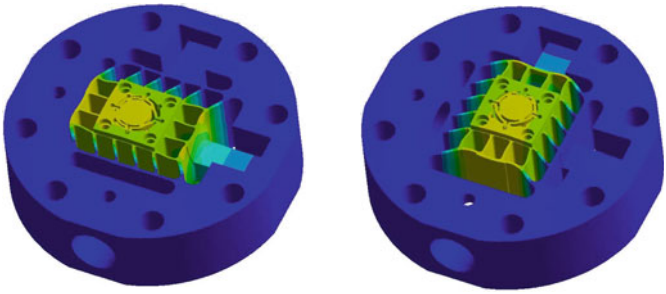


Fig. 3.20 FE simulated first mechanical resonance modes along the actuation direction of the *xy*-nanopositioner

Meshing was first generated using ANSYS "auto-mesh" tool. Meshing near the flexure was then refined manually to obtain more precise results. The material of the nanopositioner is aluminum alloy Al7075 with Young's modulus of 72 GPa. For modal analysis, motions of the base, in which the *xy*-nanopositioner was attached, were fixed during simulation. The simulated first resonant frequencies of the stage along the *x* and *y* axes both are at 12.7 kHz as illustrated in Fig. 3.20.

For stiffness analysis, a point force was applied to each actuation direction of the nanopositioner and the corresponding displacement was recorded. The stiffness is calculated as force over displacement. The simulated stiffness are 13.1 N/μm for both lateral axes.

3.5.1.5 Piezoelectric Stack Actuators

Driving piezoelectric stack actuators at high frequencies amounts to a significant challenge. It is well known that the larger the capacitance of a piezoelectric stack actuator, the larger the current required to drive the actuator at high speed [38]. The required current may exceed the current limit of the driving amplifier. Thus, the piezoelectric stack actuators must be chosen carefully to ensure their suitability for the specific system. The piezoelectric stack actuators selected to drive the stage along the *x* and *y* directions are Noliac SCMAP07. These actuators have a capacitance of 380 nF. Table 3.4 contains detailed specifications of these actuators. Three PiezoDrive PDL200 amplifiers were used to drive the piezos. The maximum operating frequency of driving a triangular signal can be estimated as $f^{3\,\mathrm{dB}} = 0.1/190\,\mathrm{C}$ [59]. With the setup described above, it is estimated that the fast lateral axis of the nanopositioner can be driven with a triangular signal up to 1.4 kHz, at the full voltage range of 200 V, without exceeding the current limit of the amplifier.

Table 3.4 Specifications of the piezoelectric stack actuators

Piezo	Dimension (mm)	Capa. (nF)	Stroke @ 200 V (μm)	Stiff. (N/μm)	Res. freq. (kHz)
X, Y	5×5×10	380	11.8	283	220

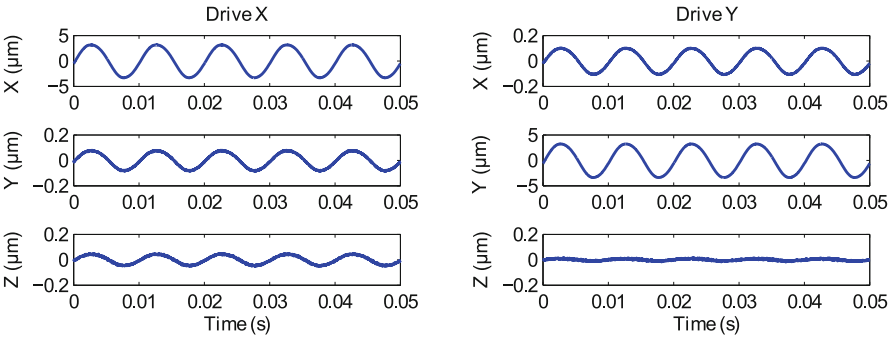


Fig. 3.21 Displacements and cross-coupling motions of the *xy*-nanopositioner

3.5.1.6 Experimental Outcomes

To evaluate the displacement range and coupling between various axes of the nanopositioner, a 100-Hz sine-wave of 200 V was applied to each of the piezoelectric stack actuators. Displacements and cross-coupling motions of the nanopositioner were measured using the ADE Technologies 8810 capacitive sensors and results are plotted in Fig. 3.21. The measured full displacement range of the x and y axes are 6.5 μm and 6.6 μm, respectively. When the x-axis is driven by the 200-V sine-wave, the induced cross-coupling motions at the y and z axes are 2.5 % (−31.8 dB) and 1.58 % (−36 dB) of the x displacement, respectively. When the y axis is driven, it induces 3.1 % (−29.9 dB) and 0.53 % (−45.5 dB) in the x and z displacement, respectively. Although these couplings are quite small, they can be made even smaller by using a feedback control technique, such as that reported in [101].

Frequency responses of the x- and y-axes of the nanopositioner were measured using a dual channel spectrum analyzer (HP 35670A). Figure 3.22 plots the frequency response functions of the two axes. The first dominant mode for the two lateral axes are 13.7 kHz and 11.6 kHz, respectively. The measured resonance frequencies are within the 10 % range of those predicted by finite-element analysis. This is most likely due to the machining errors when the nanopositioner was being manufactured, and the variations in stiffness of piezoelectric stack actuators. Nevertheless, the measured resonances are in reasonable agreement with the FE simulations.

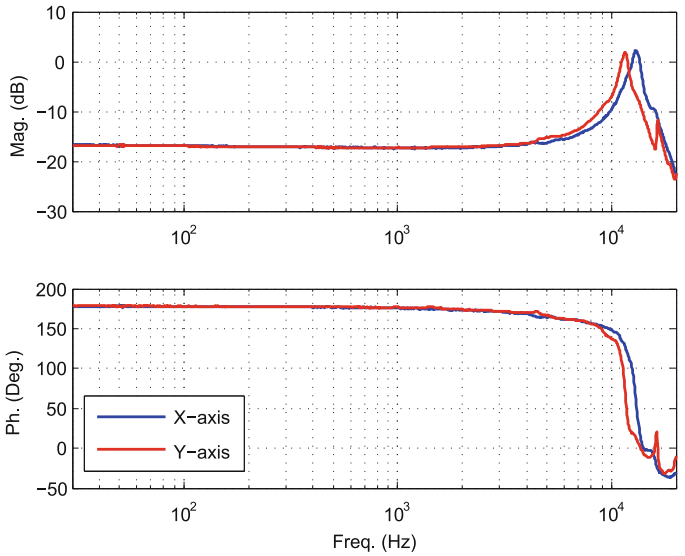


Fig. 3.22 Measured frequency responses of the x- and y-axis. The first dominant peak are at 13.7 kHz and 11.6 kHz, respectively

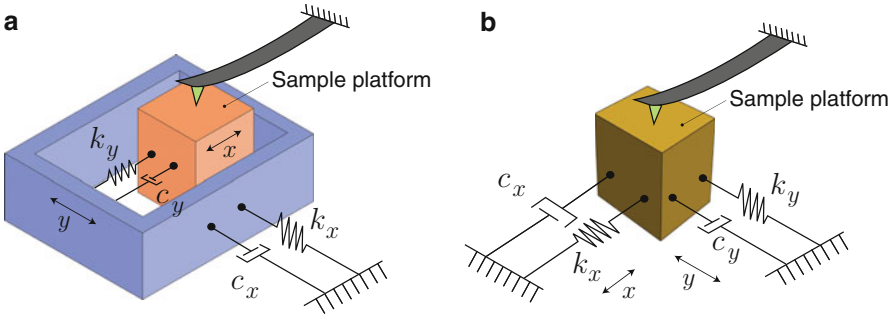


Fig. 3.23 Simplified models for two-axis scanning: (**a**) serial- and (**b**) parallel-kinematic configuration. The spring and damping constants include the effects of the piezoactuators and added flexures in each direction

3.5.2 Serial-Kinematic Nanopositioner

The design of a high-speed serial-kinematic nanopositioner is described in this section as an illustrative example. It is pointed out that for scanning in two directions, there are two basic configurations: serial- and parallel-kinematics as illustrated in Fig. 3.23. In a serial-kinematic system, for example, the design used by several commercial vendors of scanning stages and in [4, 35], there is exactly one actuator (and sensor) for each degree of freedom (see Fig. 3.23a). One disadvantage of this design is the inability to measure (and correct for) parasitic motion such

as runout or guiding error. Although the serial configuration is simple to design, a penalty is that high resonance frequency can only be achieved in one axis. But, this also implying that the high-speed axis requires high power and wide bandwidth performance, reducing overall cost compared to a parallel-kinematic design.

In the simplified model shown in Fig. 3.23, the effective stiffnesses and damping effect for both the serial- and parallel-kinematic configurations include the flexures and the actuators along each direction. To achieve high resonance frequencies, the effective stiffnesses should be as high as possible while achieving the desired range of motion. The effects of inertial force generated by the sample platform during scanning must also be taken into account. The flexures must provide enough preload to avoid exposing the stack actuator to damaging tensional forces.

In the design of a serial-kinematic nanopositioner, while the resonance frequency of the fast axis is of primary concern, the slow-axis resonance frequency can essentially be ignored. For example, the scan rate of the slow axis is one-hundredth the scan rate of the fast axis when acquiring a 100×100 pixel AFM image. Therefore, the fast-scan axis can be designed independently without any significant consideration for performance implications on the slow-scan axis. As previously stated, for scanning-type applications, one lateral axis operates much faster than the other, and thus the serial-kinematic configuration is well-suited for these types of applications [6, 38].

For the design presented below, the objective is to create a nanopositioner for high-bandwidth applications that include video-rate scanning probe microscopy. The proposed design offers approximately $9 \times 9 \times 1$ μm range of motion with kHz bandwidth. Vertically-stiff, double-hinged serial flexures are employed to guide the motion of the sample platform to minimize parasitic motion (runout) and off-axis effects (refer to example in Fig. 3.8).

3.5.2.1 Flexure Design for Lateral Positioning

The basic layout for a serial-kinematic design is shown in Fig. 3.24, where the high-speed (x-axis) stage is nested inside of the low-speed (y-axis) stage. The stage body is typically manufactured from 7075 aluminum or Invar using the wire EDM process to create a monolithic platform. The sample platform is located on the x-stage body, and vertical motion is achieved with a piezo-stack actuator embedded into the x-stage body [6]. Compliant flexures with improved vertical stiffness to minimize out-of-plane motion are used to guide the motion of the sample platform. The flexures are strategically-placed to minimize the sample platform's tendency to rotate (θ_x, θ_y, θ_z) at high frequencies. Also, the stage is designed to ensure that the first resonance in all three axes are axial (piston) modes, rather than off-axis modes which can severely limit performance.

For translational motion, u_i ($i = x,\ y,\ z$), the single DOF mechanical resonance is given by $f_{u_i,0} = \frac{1}{2\pi}\sqrt{\frac{k_i}{m_i}}$, where m_i and k_i are the effective translational mass and stiffness, respectively. Likewise for rotational motion, θ_i ($i = x,\ y,\ z$), the first resonance is $f_{\theta_i,0} = \frac{1}{2\pi}\sqrt{\frac{k_{\theta_i}}{J_i}}$, where J_i and k_{θ_i} are the effective mass moment

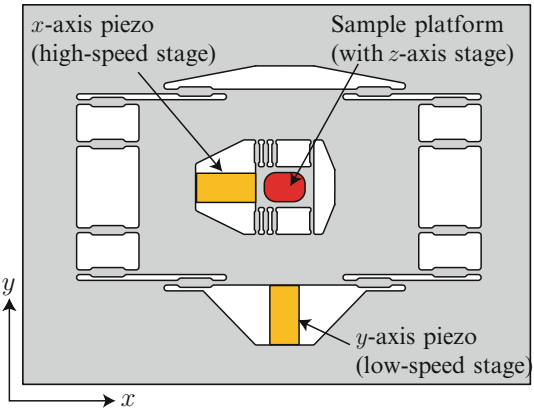


Fig. 3.24 A serial-kinematic nanopositioner (*top view*), showing the lateral stages and details of the sample platform and z-stage

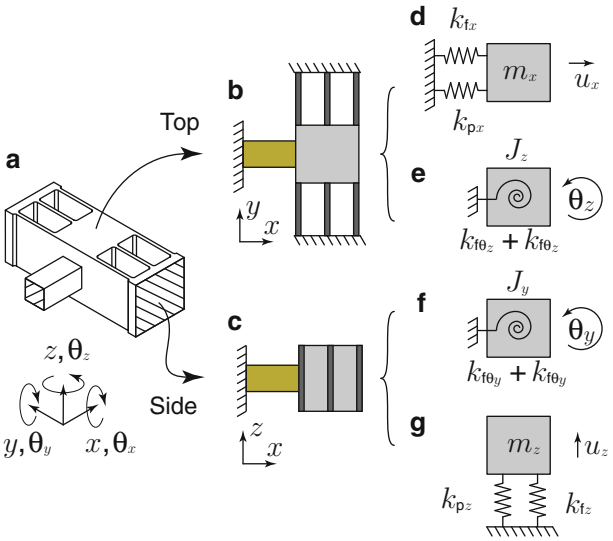


Fig. 3.25 Generic flexure-guided stage simplified to single degree-of-freedom systems modeling the dominant modes

of inertia and rotational stiffness, respectively. To insure that actuation modes occur before the out-of-plane modes, the strategy taken is to optimize the stage geometry and flexure configuration so that the out-of-plane stiffness-to-mass ratios (k_z/m_z, $k_{\theta y}/J_y$, $k_{\theta z}/J_z$) are higher than the actuation stiffness-to-mass ratio k_x/m_x. Figure 3.25 shows the simplification of a high-speed x-stage into single DOF systems to model four of the dominating resonance modes. The top and side views are broken down to show the effective springs and masses affecting the body for (d) actuation u_x, (e) and (f) rotation θ_z and θ_y, and (g) vertical u_z modes. Damping is omitted for convenience.

The vertical stiffness of the x- and y-stages is increased by (1) increasing the number of flexures, (2) utilizing shorter (effective length) flexures, and (3) converting the flexures from constant rectangular cross-section beam flexures to a serial-compliant double-hinged flexure with a "rigid" center connecting link (see Fig. 3.8).

The stiffness of a flexure is defined as the ratio of a load F and the resulting displacement u. The displacements and loads are: translational displacement u_i, rotational displacement θ_i, translational force F_i acting on a point in the i direction, and moment M_i (torque T) acting about the i axis (θ_i), respectively, where $i = x, y, z$. Figure 3.6 illustrates the corresponding directions of the displacements and loads acting on the free end of a fixed/free cantilever beam which models a beam flexure. The in- and out-of-plane compliances for a fixed/free beam is derived using Castigliano's second theorem [16, 44, 84]. The compliance equations are then used to derive equations for the actuation and vertical stiffness k_i of a fixed/roller guided beam shown in Fig. 3.26(a1)–(a3). It is pointed out that the fillet radius is considerably smaller compared to the flexure length and therefore has minimal effect on the flexure stiffness. For this reason, to simplify the flexure stiffness equations in this initial analysis, the compliance equations are derived for a beam with a constant cross-sectional thickness.

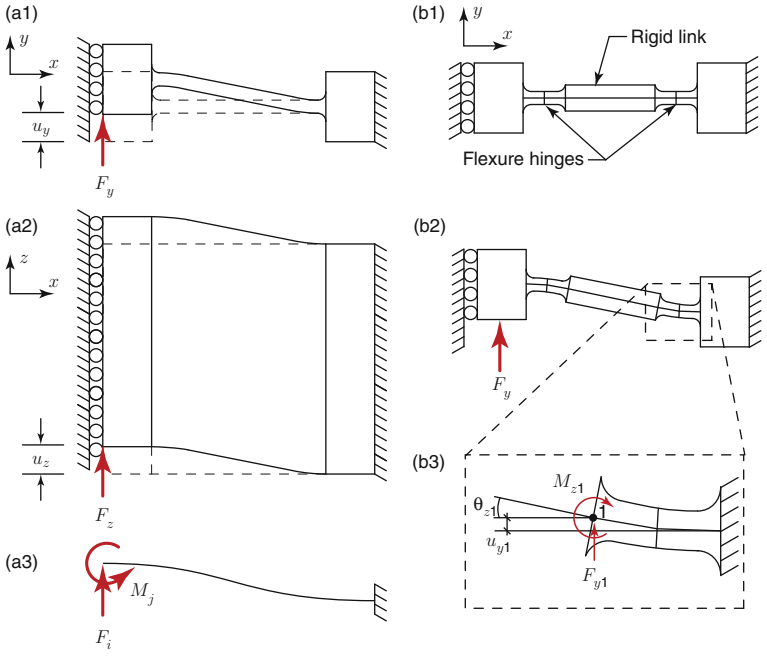


Fig. 3.26 Corner-filleted and center-thickened flexures showing loads and deformations: (**a1**) *top* and (**a3**) *side views* showing displacement caused by force F_i, for $i = x, y$, in a fixed/guided end configuration, and (**a3**) loads acting on the free end of a fixed/free beam for a corner-filleted. (**b1**) *top view*, (**b2**) *top view* with applied load, and (**b3**) expanded view of corner-filleted flexure hinge

For a fixed/free beam of rectangular cross section the total strain energy is

$$\begin{aligned} U &= U_{\text{axial}} + U_{\text{torsion}} + U_{\text{bending}} + U_{\text{shear}} \\ &= \int_0^L \left[\frac{F^2}{2AE} + \frac{T^2}{2GJ} + \frac{M^2}{2EI} + \frac{\alpha V^2}{2GA} \right] dx, \end{aligned} \tag{3.65}$$

where L is the beam length, A is the cross-sectional area of the beam, h is the height, t is the thickness, E is the Young's modulus, $G = \frac{E}{2(1+\nu)}$ is the shear modulus, ν is the Poisson's ratio, $J = ht^3 \left[\frac{1}{3} - 0.21 \frac{t}{h} \left(1 - \frac{t^4}{12h^4} \right) \right]$ is the approximate torsional moment of inertia [107], $I = \frac{ht^3}{12}$ is the second moment of inertia about the vertical z axis, V is the shear force, and α is the shape factor for the cross section used in the shear equation (for a rectangular cross section $\alpha = 6/5$) [16, 60, 107].

Applying Castigliano's second theorem, the displacement of a point in a given direction u_i, θ_i is the partial derivative of the total strain energy with respect to the applied force, i.e.:

$$u_i = \frac{\partial U}{\partial F_i}; \qquad \theta_i = \frac{\partial U}{\partial M_i}. \tag{3.66}$$

From here the compliance is simply found by dividing the displacement by the applied load, i.e.:

$$C_{u_i,F_j} = \frac{u_i}{F_j}; \qquad C_{\theta_i,M_j} = \frac{\theta_i}{M_j}. \tag{3.67}$$

For example, the compliance of the rectangular cross section fixed-free beam in Fig. 3.6 due to a point load in the y direction starts with the total strain energy

$$U = \int_0^L \frac{M(x)^2}{2EI(x)}\, dx + \int_0^L \frac{\alpha V(x)^2}{2GA(x)}\, dx, \tag{3.68}$$

where $A(x)$ and $I(x)$ are constant. The coordinate system is placed on the free end of the flexure as shown in [44] where the shear is $V(x) = F_y$ and moment is $M(x) = F_y x$. The total strain energy for the applied load is

$$U = \frac{F_y^2}{2EI} \int_0^L x^2\, dx + \frac{\alpha F_y^2}{2GA} \int_0^L dx = \frac{F_y^2 L^3}{6EI} + \frac{\alpha F_y^2 L}{2GA}. \tag{3.69}$$

Therefore, the resultant displacement is

$$u_y = \frac{\partial U}{\partial F_y} = \frac{F_y L^3}{3EI} + \frac{\alpha L F_y}{GA}, \tag{3.70}$$

and the compliance is

$$C_{22} = \frac{u_y}{F_y} = \frac{L^3}{3EI} + \frac{\alpha L}{GA}. \tag{3.71}$$

The compliances are then used to form the compliance matrix $\mathbf{C}$ which is defined as the ratio of the displacement $\mathbf{U} = \begin{bmatrix} x\ y\ \theta_z\ z\ \theta_y\ \theta_x \end{bmatrix}^T$ for a given load $\mathbf{L} = \begin{bmatrix} F_x\ F_y\ M_z\ F_z\ M_y\ M_x \end{bmatrix}^T$, hence the displacement vector is

$$\begin{Bmatrix} u_x \\ u_y \\ \theta_z \\ u_z \\ \theta_y \\ \theta_x \end{Bmatrix} = \begin{bmatrix} C_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & C_{22} & C_{23} & 0 & 0 & 0 \\ 0 & C_{23} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & C_{45} & 0 \\ 0 & 0 & 0 & C_{45} & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix} \begin{Bmatrix} F_x \\ F_y \\ M_z \\ F_z \\ M_y \\ M_x \end{Bmatrix}. \tag{3.72}$$

For a constant cross section fixed/free beam the compliances are $C_{11} = \frac{L}{AE}$, $C_{22} = \frac{L^3}{3EI} + \frac{\alpha L}{GA}$, $C_{23} = \frac{L^2}{2EI}$, $C_{33} = \frac{L}{EI}$, $C_{44} = \frac{4L^3}{Eh^3t} + \frac{\alpha L}{GA}$, $C_{45} = \frac{6L^2}{Eh^3t}$,$C_{55} = \frac{12L}{Eh^3t}$, and $C_{66} = \frac{L}{GJ}$. For a long slender beam, shear strain has little effect and therefore can be ignored in C_{22}. For a short beam with a significant height-to-length aspect ratio, such as the vertical displacement of the flexure shown in Fig. 3.26(a2), much of the deflection is in shear, and therefore it can not be ignored.

The displacement vector equation presented above is used to solve for the actuation stiffness k_y and vertical stiffness k_z of a fixed/guided flexure beam, i.e., $F_i/u_i = k_i$. Torsional stiffness is not investigated because the θ_x rotational mode is largely dependant upon the vertical flexure stiffness when the flexures are placed at the corners of the stage body. Figure 3.26(a3) shows the applied load and the expected deflection curve of the flexure in both the (a1) actuation direction and (a2) vertical direction. The active load being applied to the flexure is the in-plane force F_i. The resultant moment $M_j = -F_iL/2$ is caused by the roller-guided end constraint. Therefore, the flexure displacement in the actuation direction u_y due to the applied force F_y and moment $M_z = -F_yL/2$ is

$$\begin{aligned} u_y &= C_{22}F_y + C_{23}M_z = C_{22}F_y - C_{23}F_yL/2 \\ &= F_y\left[\frac{L^3}{3EI} + \frac{\alpha L}{Ght} - \frac{L}{2}\frac{L^2}{2EI}\right]. \end{aligned} \tag{3.73}$$

Taking the ratio of the applied load to the displacement, the actuation stiffness (neglecting shear) is

$$k_y = \frac{F_y}{y} = \left[\frac{L^3}{12EI} + \frac{\alpha L}{Ght}\right]^{-1} \cong \frac{12EI}{L^3}. \tag{3.74}$$

Using the same method, the displacement of the flexure in the vertical direction u_z is

$$u_z = C_{44}F_z + C_{45}M_y = C_{44}F_z - C_{45}F_zL/2$$
$$= F_z\left[\frac{4L^3}{Eh^3t} + \frac{\alpha L}{Ght} - \frac{L}{2}\frac{6L^2}{Eh^3t}\right]. \tag{3.75}$$

Similarly, the vertical stiffness is

$$k_z = \left[\frac{L^3}{Eh^3t} + \frac{\alpha L}{Ght}\right]^{-1}. \tag{3.76}$$

Because of the high aspect ratio in the vertical direction, shear can not be ignored. It is pointed out that Eqs. (3.74) and (3.76) are used to design the flexures to meet certain specifications, such as lateral and vertical stiffness.

Overall, the effective vertical stiffness can be improved to increase the out-of-plane stiffness by (1) increasing the number of flexures n, (2) decreasing the flexure length L, and (3) thickening the center section of a beam flexure to create a serial-compliant double-hinged flexure. Additionally, flexure placement is important to help increase rotational stiffness. Increasing the length (and width) of a stage and placing flexures at the corners of the moving platform increase rotational stiffness of the platform. However, the cost of increasing the size of the platform is increasing overall mass, thus lowering the mechanical resonance.

3.5.2.2 Results: Comparison Between FEA and Experimental Results

The first three modes for the x and y stages were predicted using the *frequency* tool in COSMOSWorks (FEA). It is assumed that the resonances of the y-stage would not be excited by the dynamic motion of the inner nested x-stage. This allows the design shown in Fig. 3.27a to be broken down into the low-speed y-stage [Fig. 3.27(b1)–(b3)] and high-speed x-stage [Fig. 3.27(c1)–(c3)]. The boundary faces of each stage (shown hatched) have a fixed boundary condition. All contacting components are bonded together with compatible mesh. The meshing is done at "high quality" with refined meshing at the flexure fillets and pivot points (0.25 mm minimum element size on surfaces). The materials used and their corresponding mechanical properties are as follows: aluminum: $E = 72\,\text{GPa}$, $\nu = 0.33$, $\rho = 2700\,\text{kg/m}^3$; steel: $E = 200\,\text{GPa}$, $\nu = 0.28$, $\rho = 7800\,\text{kg/m}^3$; piezo-stack: $E = 33.9\,\text{GPa}$, $\nu = 0.30$, $\rho = 8000\,\text{kg/m}^3$; and alumina: $E = 300\,\text{GPa}$, $\nu = 0.21$, $\rho = 3960\,\text{kg/m}^3$, where the modulus for the piezo-stack was calculated from the stiffness and blocking force. The predicted first mechanical resonance for the y- and x-stage are 5.96 kHz and 25.9 kHz, respectively, all of which are in the corresponding stage actuation direction as preferred. Simulated FEA frequency response is done using the *Linear Dynamic (Harmonic)* tool in COSMOSWorks. A constant amplitude sinusoidal force is applied in the actuation direction at the corners of the piezoactuator/stage

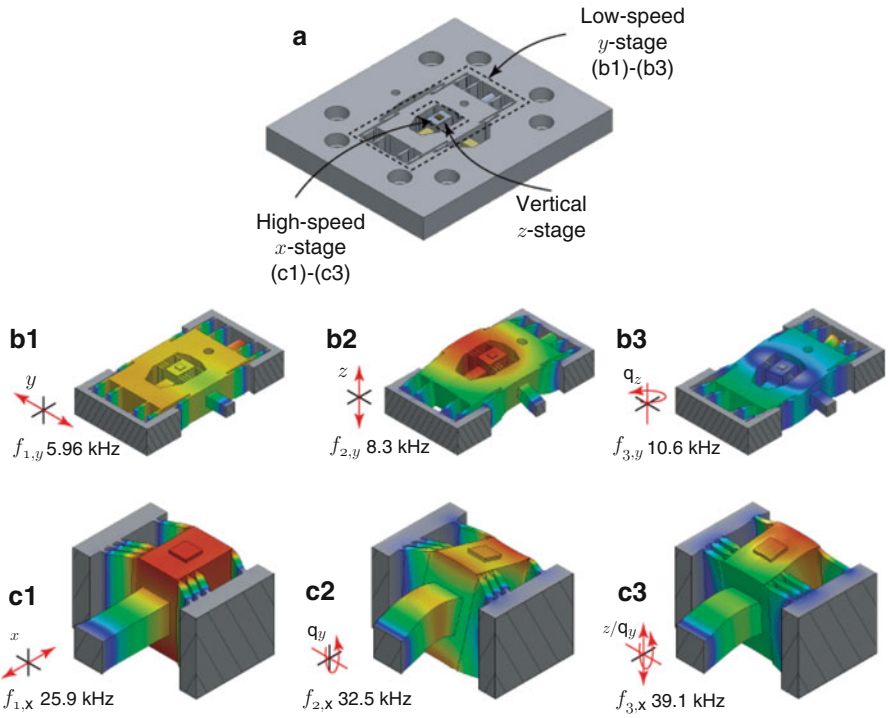


Fig. 3.27 Finite-element analysis results showing first four modes: (**a**) high-speed scanning stage; (**b1**)–(**b3**) low-speed *y*-stage; (**c1**)–(**c3**) high-speed *x*-stage. Each stage section is designed to have the first mechanical resonance to occur in the actuation direction

interfaces. The force generated is assumed proportional to the applied voltage. A global modal damping ratio of 0.025 is applied to simulate the damping of aluminum alloy and to produce a gain of 20 dB. Figure 3.28(a1), (b1) show the predicted frequency response plots for the *x*- and *y*-axes with the resonant peaks occurring at 25.9 kHz and 5.96 kHz, respectively. Overall, the FEA results and the experimental measurements show good agreement.

3.6 Design Example of a Vertical Nanopositioner

There are increasing demands for vertical nanopositioning systems in applications such as high-speed vertical stage for scanning probe microscopy [23, 34, 98], objective focusing stages [24, 77], nanofabrication [85, 86], mask, and wafer surface inspection [12]. Typically, flexures are used to provide preloads to piezoelectric stack-driven nanopositioners [100, 105], and to minimize off-axis motions from the actuation direction [34, 98].

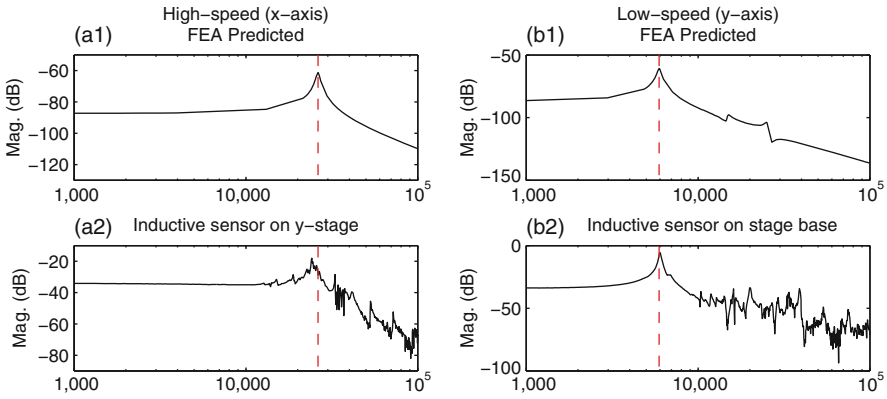


Fig. 3.28 (**a1**)–(**a2**) Comparison of predicted and measured frequency response functions for the high-speed stage (*x*-axis), (**b1**)–(**b2**) the low-speed stage (*y*-axis). The *vertical dashed line* is used to compare the experimentally measured results to the FEA predicted first resonance peak

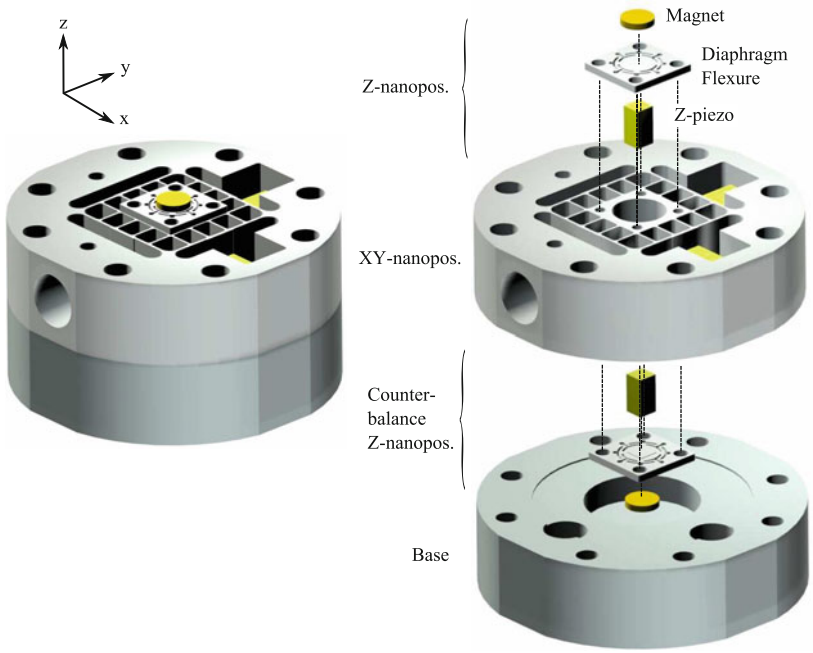


Fig. 3.29 Assembled and exploded views of the *z*-nanopositioner design

A design example of a high-speed vertical nanopositioner [98] is presented here. The *z*-stage is mounted on the *xy*-stage presented in Sect. 3.5.1 (see Fig. 3.29) to achieve the three translational DOFs (degree-of-freedoms) for high-speed atomic force microscopy. The vertical nanopositioner utilizes a diaphragm flexure

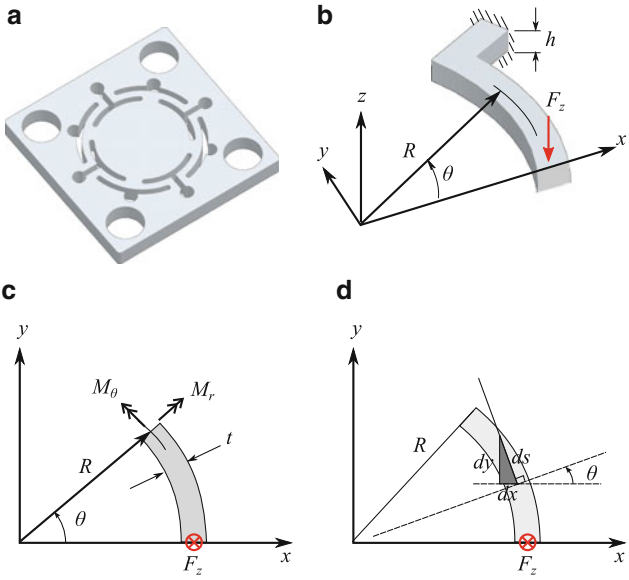


Fig. 3.30 Diaphragm flexure of the z-nanopositioner. (**a**) Diaphragm flexure. (**b**) Curved-beam model. (**c**) Cross-sectional bending and torsional moments of the beam. (**d**) Beam element

in Fig. 3.30a to guide and preload the z-axis piezoelectric stack actuator. The diaphragm flexure is constructed from eight curved beams. The curved beams are fixed at one end and guided at the other. To estimate the effective stiffness of the diaphragm flexure, a point load F_z is applied to the guided end. The total strain energy of the curved beam flexure (neglecting shearing) is

$$U = \int_0^L \frac{M_r^2}{2EI} ds + \int_0^L \frac{M_\theta^2}{2GJ} ds, \tag{3.77}$$

where M_r and M_θ as shown in Fig. 3.30c are the cross-sectional bending and torsional moments, which vary with θ, E is the Young's modulus, $G = E/2\,(1+\upsilon)$ is the shear modulus, υ is the Poisson's ratio, $I = h^3 t/12$ is the second moment of inertia, $J = \left(ht^3 + h^3 t\right)/12$ is the polar moment of inertia, ds is the length of a beam element, and L is the length of the beam. According to the Castigliano's second theorem, the deflection δ_z of the curved beam flexure can be calculated by taking the partial derivative of U with respect to F_z, that is

$$\delta_z = \frac{\partial U}{\partial F_z} = \frac{1}{EI} \int_0^L M_r \frac{\partial M_r}{\partial F_z} ds + \frac{1}{GJ} \int_0^L M_\theta \frac{\partial M_\theta}{\partial F_z} ds. \tag{3.78}$$

The equilibrium equations for moment along the x and y directions are [18]

$$M_r \cos\theta - M_\theta \sin\theta + F_z y = 0, \tag{3.79}$$

$$M_r \sin\theta + M_\theta \cos\theta + F_z (R - x) = 0. \tag{3.80}$$

Solving for M_r and M_θ,

$$M_r = -F_z y \cos\theta - F_z (R - x) \sin\theta, \tag{3.81}$$

$$M_\theta = F_z y \sin\theta - F_z (R - x) \cos\theta. \tag{3.82}$$

The following procedure is employed in order to perform the integration of Eq. (3.78) over the variable x in the Cartesian coordinates [18]. The curve of the beam is a circular profile, therefore:

$$x^2 + y^2 = R^2, \tag{3.83}$$

$$y = \sqrt{R^2 - x^2}, \tag{3.84}$$

where R is the radius of curvature of the beam. By differentiating Eq. (3.84), we have

$$\frac{dy}{dx} = -\frac{x}{\sqrt{R^2 - x^2}}. \tag{3.85}$$

As shown in Fig. 3.30c, the length of a beam element ds is

$$ds = \sqrt{dx^2 + dy^2} = dx\sqrt{1 + \left(\frac{dy}{dx}\right)^2}. \tag{3.86}$$

And $\sin(\theta)$ and $\cos(\theta)$ are expressed as

$$\sin\theta = \frac{dx}{ds} = \frac{1}{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}, \tag{3.87}$$

$$\cos\theta = \frac{dy}{ds} = \frac{-dy/dx}{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}. \tag{3.88}$$

By substituting Eqs. (3.81), (3.82), (3.86)–(3.88) and $\partial M_r/\partial F_z$, $\partial M_\theta/\partial F_z$ into Eq. (3.78), and performing integration from $x = R\cos(\pi/4)$ (where $\pi/4$ is the angle of the beam) to $x = R$, we have

$$\delta_z = F_z R^2 \left(\frac{T_1}{EI} + \frac{T_2}{GJ} \right), \tag{3.89}$$

where T_1 and T_2 are

$$T_1 = -\int_{\frac{R}{\sqrt{2}}}^{R} \frac{1}{\sqrt{\frac{R^2}{R^2 - x^2}}} dx, \tag{3.90}$$

$$T_2 = -\int_{\frac{R}{\sqrt{2}}}^{R} \frac{R - x}{(R + x)\sqrt{\frac{R^2}{R^2 - x^2}}} dx. \tag{3.91}$$

From Eq. (3.89), the stiffness of the curved beam is $K_z = F_z/\delta_z$. The diaphragm flexure consists of eight curved beams arranged in parallel. Therefore, the total stiffness of the flexure is

$$K_f = \frac{8F_z}{\delta_z}. \tag{3.92}$$

For the chosen flexure dimensions and material as documented in Table 3.5, K_f is calculated to be 5.17 N/μm.

3.6.1 Resonance Frequency

Resonance frequency of the single-mounted z-nanopositioner can be estimated by modeling it as a two DOF spring-mass-damper system [69] as shown in Fig. 3.31. In this model, the upper half of the piezoelectric stack actuator is lumped to the effective mass of the flexure and sample, the lower half of the actuator is lumped to the effective mass of the support structure, that is the xy-nanopositioner. Also, c_1 and c_2 represent the effective damping coefficients, and k_1 and k_2 are the effective stiffnesses. We have $m_1 = m_p/2 + m_s$, $m_2 = m_p/2 + m_f$, $c_1 = c_s$, $c_2 = c_p + c_f$, $k_1 = k_s$, and $k_2 = k_p + k_f$. The subscripts p, f, and s stand for piezoelectric stack actuator, diaphragm flexure, and support structure, respectively. By using the lumped spring-mass-damper system, two resonances (f_1 and f_3) and one anti-resonance (f_2) can be obtained [69], as

$$f_1 \approx \frac{1}{2\pi} \sqrt{\frac{k_1}{m_1 + m_2}};$$

Table 3.5 Dimensions and material properties of flexure and piezoelectric stack actuator

Diaphragm flexure	Values
Radius of curvature, R (mm)	2.99
Thickness, t (mm)	0.38
Height, h (mm)	1.1
Effective mass, m_f (g)	0.32
Stiffness, k_f (N/μm)	5.17[a]
Young's modulus, E (GPa)	72
Poisson's ratio, υ	0.33
Density, ρ (kg/m^3)	2770
Piezo	Values
Dimension (mm)	3 × 3 × 6
Mass, m_p (g)	0.41
Stiffness, k_p (N/μm)	65
Young's modulus, E_p (GPa)	45
Poisson's ratio, υ_p	0.3
Capacitance, C_p (nF)	70
Travel range @ 200 V (μm)	5.5
Support structure (xy-stage)	Values
Mass, m_s (g)	5.72
Vertical stiffness, k_s (N/μm)	81[b]
Density, ρ (kg/m^3)	7500

[a]Obtained using Eq. (3.92)
[b]Obtained using FE simulations

$$f_2 \approx \frac{1}{2\pi}\sqrt{\frac{k_1}{m_1}};$$
$$f_3 \approx \frac{1}{2\pi}\sqrt{\frac{(m_1+m_2)\,k_2}{m_1 m_2}}. \tag{3.93}$$

Here, f_1 is the vertical resonant mode of the xy-nanopositioner, and f_3 is the resonant mode of the preloaded piezoelectric stack actuator with flexures. The calculated resonance frequencies f_1 and f_3 of the single-mounted z-nanopositioner design are 17.8 kHz and 61 kHz, respectively.

3.6.2 Finite-Element-Analysis

ANSYS was used to perform finite-element (FE) analysis of the z-nanopositioner. The flexure is deformed using a piezoelectric stack actuator that generates the requisite vertical displacement. A magnet is glued onto the flexure as illustrated in Fig. 3.32. The magnet's function is to hold a metal sample holder. It has a diameter

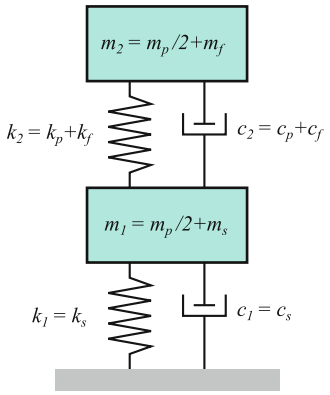


Fig. 3.31 A 2-DOF spring-mass model of the z-nanopositioner where m, c, and k are mass, damping coefficient, and stiffness, respectively. The subscripts p, f, and s are referred to the piezoelectric stack actuator, diaphragm flexure, and support structure (the xy-nanopositioner) respectively

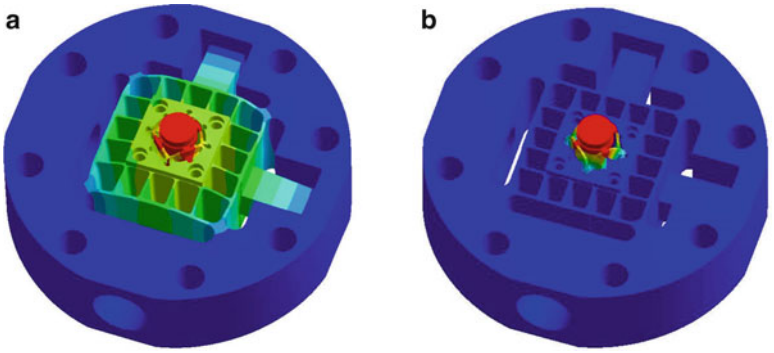


Fig. 3.32 (Finite-element simulations. (**a**) First vertical resonant mode of the xy-nanopositioner: 19.7 kHz. (**b**) First vertical resonant mode of the z-nanopositioner: 65 kHz

of 5 mm and a thickness of 1 mm. The flexure is made of aluminum alloy (Al7075). The piezoelectric actuator is a Noliac SCMAP06 stack actuator. Dimensions and material properties of the actuator, which were used to construct the ANSYS model, can be found in Table 3.5.

ANSYS simulations predicted that the first vertical resonance frequency of the xy-nanopositioner is at 19.7 kHz. The vertical resonant mode of the z-nanopositioner appears at 65 kHz (see Fig. 3.32). The simulated stiffness along the z-axis is 5.15 N/μm. The analytical resonance frequencies and stiffness are in reasonable agreement with the FE simulations.

3.6.3 Travel Range

To estimate travel range of the nanopositioner, we may use the formulation $\Delta L = \Delta L_o \cdot k_p / \left(k_f + k_p\right)$ [83, 100], where ΔL is the displacement with external spring load, ΔL_o is the maximum nominal displacement of the piezoelectric

stack actuator without an external spring load, k_f is the flexure stiffness, and k_p is the piezoelectric stack actuator stiffness. The estimated displacement of the z-nanopositioner is 5.1 μm.

3.6.4 Inertial Cancellation

The z-nanopositioner is mounted to a parallel-kinematics xy-nanopositioner presented in Sect. 3.5.1 and illustrated in Fig. 3.29. When the z-nanopositioner is only comprised of a single-nanopositioner, as is the case in [97, 103] (and illustrated in Fig. 3.33a), the fast movements of the piezoelectric stack actuator tend to induce impulsive forces which propagate through, and excite the resonant modes of the xy-nanopositioner [5, 104]. This resonant mode, which often lives at a lower frequency than that of the vertical axis, in turn, causes vibrations in the z-nanopositioner. The frequency response of the single-nanopositioner configuration shows that the vertical resonant mode of the xy-nanopositioner is the dominant mode with a phase drop of more than 45°. This mode may need to be suppressed by means of feedback control to avoid vibrations. In this configuration, bandwidth of the vertical feedback loop is limited by the low resonance frequency of the xy-stage.

Equation (3.93) shows that the vertical resonant mode of the xy-stage, f_1, can be increased by increasing the vertical effective stiffness k_s of the xy-nanopositioner. This can be achieved by adding more beam flexures to the xy-nanopositioner. However, these flexures will increase the effective mass of the device which in turn will reduce the resonance frequency. Furthermore, the additional flexures result in a reduction in the xy travel range. To avoid this, a dual-mounted z-nanopositioner configuration, illustrated in Fig. 3.33b, is used to counterbalance the impulsive forces of the z-actuator [34, 35, 98]. This inertial counterbalance method avoids the need to change k_s through mechanical design. The bottom z-actuator cancels the impulsive force exerted by the top actuator which, when the mass is balanced properly, eliminates the first vertical resonant mode. Other inertial counterbalance techniques can also be found in [5].

The effectiveness of the dual-mounted nanopositioner is sensitive to the weight difference between the opposing piezoelectric stack actuators. This is illustrated in Figs. 3.33(b2), (b3). If the sample and the dummy weights are perfectly balanced, the vertical mode of the xy-nanopositioner will not be triggered. In this case, the dominant mode is the vertical mode of the z-nanopositioner where, by design, is at least 4 times higher than that of the xy-nanopositioner. Thus, in this case, bandwidth of the vertical feedback loop is limited by the resonant mode of the z-nanopositioner. However, in practice, it is rather difficult to perfectly balance the weights. It is more likely that the dummy sample is slightly lighter or heavier than that of the sample sitting on top of the z-nanopositioner (see Fig. 3.33b). Nevertheless, if the weights are well-balanced, the effect of the xy-stage's dynamics on the vertical positioning stage will be negligible, possibly consisting of a very small peak in magnitude and a phase variation of less than 45°. In such a case, the dynamic mode of the xy-stage will not be the dominant mode.

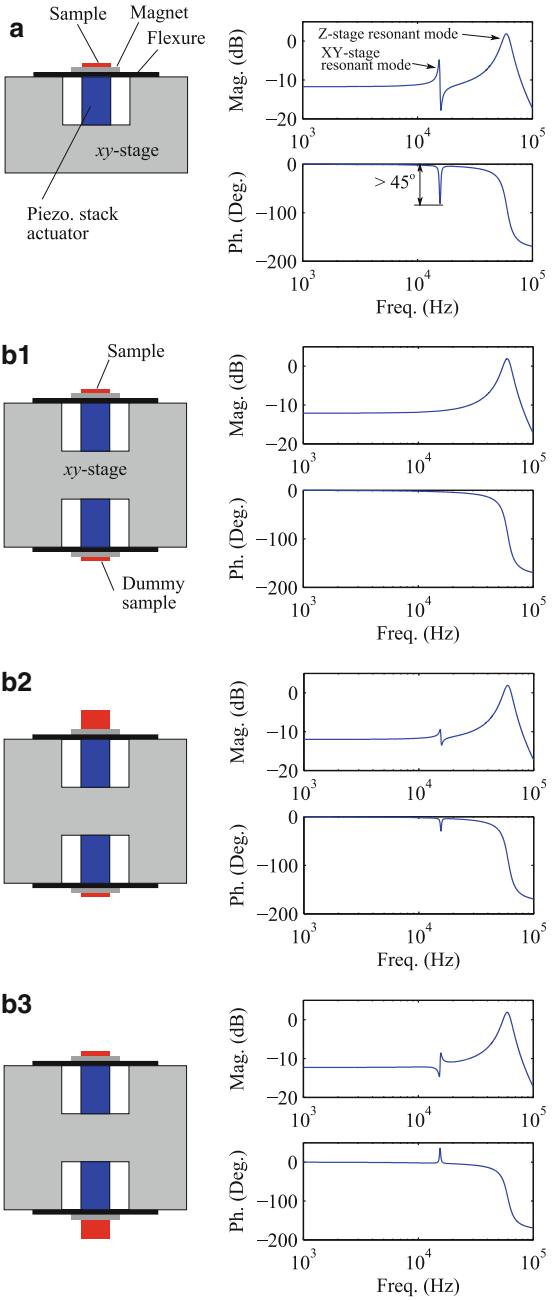


Fig. 3.33 Frequency responses of different z-nanopositioner configurations. (**a**) Single-nanopositioner configuration without inertial cancellation. (**b**) Dual-mounted z-nanopositioner configuration: (**b1**) Weights are perfectly balanced. (**b2**) Weight of the sample is slightly heavier than that of the dummy. (**b3**) Weight of the dummy sample is slightly heavier than that of the sample

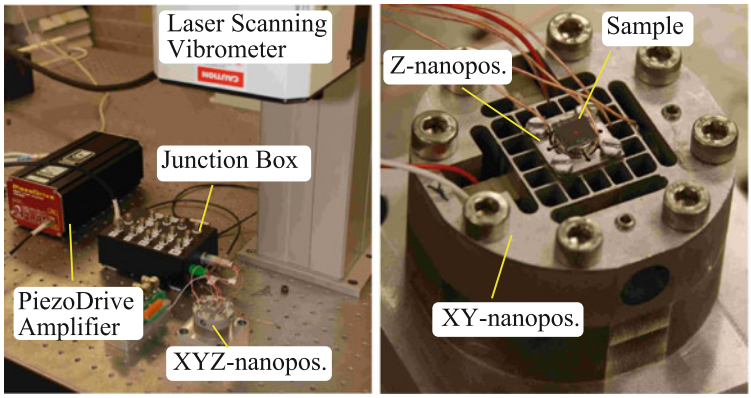


Fig. 3.34 Experimental setup of the nanopositioner. The *z*-nanopositioner is mounted to the *xy*-nanopositioner

3.6.5 Experimental Results

The experimental setup is shown in Fig. 3.34. A Polytec PSV-300 laser scanning vibrometer was used to measure the *z*-axis frequency response and travel range of the nanopositioner. A 100-Hz, 200-V peak-to-peak sinusoidal signal, which was amplified by a PiezoDrive PDL200 amplifier, was applied to drive the piezoelectric stack actuator, and the measured full-range displacement of the *z*-nanopositioner is 5 μm. The displacement is plotted in Fig. 3.35.

A pseudo-random input signal with a frequency in the range of 1–100 kHz was applied to the piezoelectric stack actuator. The measured frequency response of the *z*-nanopositioner are plotted in Fig. 3.35a. The weight of the dummy sample is 0.03 g, which is approximately the weight of the sample. The frequency response of the *z*-nanopositioner in Fig. 3.35a shows a pole-zero pattern, indicating that the dummy mass is slightly lighter than the sample. The measured first vertical resonant mode of the single-mounted *z*-nanopositioner is 17.1 kHz. Due to the additional weights of the dual-mounted *z*-nanopositioner, the vertical resonant mode of the *xy*-nanopositioner drops from 17.1 to 15 kHz. However, its resonant peak has a magnitude of only 1.6 dB and a phase drop of less than 15°. Due to its negligible profile, this mode will not be the dominant mode that dictates the vertical feedback bandwidth. As shown in the measured frequency response, the dominant mode in the *z*-axis is the 60-kHz resonant peak. As a result, a controller may be designed to suppress the 60-kHz resonant peak in order to obtain an even higher bandwidth.

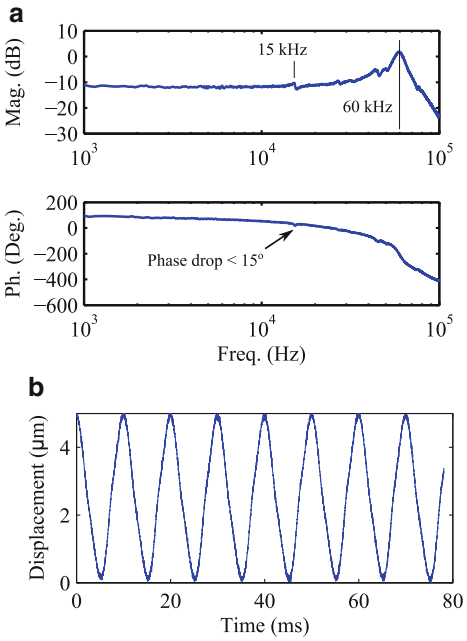


Fig. 3.35 (**a**) Measured frequency response. (**b**) Travel range of the nanopositioner

3.7 Conclusions

This chapter described the design process for high-speed flexure-guided nanopositioning stages with a focus on atomic force microscopy applications. It discussed important design considerations which include balancing trade-off between range and speed (resonance frequency), taking careful consideration of flexure designs, actuator properties and drive electronics, material selections, manufacturing techniques, and using FEA tools to optimize mechanical resonances. Design examples of serial- and parallel-kinematic nanopositioning stages were presented. A design example of a vertical nanopositioning stage with inertial counterbalance technique, which passively suppresses the low vertical resonant mode of the system was also presented.

Acknowledgements Y. K. Yong would like to acknowledge Prof. Reza Moheimani for many years of support, guidance, and valuable feedback, her colleagues at the Laboratory for Dynamics and Control of Nanosystems for providing meaningful discussions and creating an incredible research environment, and the University of Newcastle for their commitment to a young researcher. She also wishes to thank the Australian Research Council for supporting her research work. K.K. Leang acknowledges A.J. Fleming for many years of collaboration and for providing the PiezoDrive amplifiers used in the experiments. Additionally, he thanks the National Science Foundation for supporting his research work.

References

1. N.M. Abbas, D.G. Solomon, M.F. Bahari, A review on current research trends in electrical discharge machining (EDM). Int. J. Mach. Tool. Manuf. **47**, 1214–1228 (2007)
2. C. Acar, A. Shkel, *MEMS Vibratory Gyroscopes: Structural Approaches to Improve Robustness*, 1st edn. (Springer, New York, 2009). doi:10.1007/978-0-387-09536-3
3. T. Ando, N. Kodera, E. Takai, D. Maruyama, S. Kiwamu, A. Toda, A high-speed atomic force microscope for studying biological macromolecules. Proc. Natl. Acad. Sci. U. S. A. **98**, 12468–12472 (2001)
4. T. Ando, N. Kodera, T. Uchihashi, A. Miyagi, R. Nakakita, H. Yamashita, K. Matada, High-speed atomic force microscope for capturing dynamic behavior of protein molecules at work. e-J. Surf. Sci. Nanotechnol. **3**, 384–392 (2005)
5. T. Ando, T. Uchihashi, T. Fukuma, High-speed atomic force microscopy for nano-visualization of dynamic biomolecular processes. Prog. Surf. Sci. **83**(7–9), 337–437 (2008)
6. T. Ando, T. Uchihashi, N. Kodera, D. Yamamoto, A. Miyagi, M. Taniguchi, H. Yamashita, High-speed AFM and nano-visualization of biomolecular processes. Pflügers Arch. Eur. J. Physiol. **456**(1), 211–225 (2008)
7. APC International Ltd: *Piezo-Mechanics: An Introduction*. APC International Ltd., Pleasant Gap (2003)
8. A. Bazaei, Y.K. Yong, S.O.R. Moheimani, High-speed Lissajous-scan atomic force microscopy: scan pattern planning and control design issues. Rev. Sci. Instrum. **83**(6), 063701 (10pp.) (2012)
9. F. Beer, E.R. Johnston, *Mechanics of Materials* (McGraw-Hill, New York, 1992)
10. G. Binnig, C.F. Quate, Atomic force microscope. Phys. Rev. Lett. **56**(9), 930–933 (1986)
11. G. Binnig, H. Rohrer, The scanning tunneling microscope. Sci. Am. **253**, 50–56 (1986)
12. G. Borionetti, A. Bazzali, R. Orizio, Atomic force microscopy: a powerful tool for surface defect and morphology inspection in semiconductor industry. Eur. Phys. J. Appl. Phys. **27**, 101–106 (2004). doi:10.1051/epjap:2004129
13. K.H.J. Buschow, F.R. de Boer, *Physics of Magnetism and Magnetic Materials*, (Kluwer Academic Publishers, New York, 2004)
14. W.D. Callister, *Materials Science and Engineering: An Introduction* (Wiley, New York, 1994)
15. F. Cardarelli, *Materials Handbook: A Concise Desktop Reference*, 2nd edn. (Springer, London, 2000)
16. R.R. Craig, *Mechanics of Materials*, 2nd edn. (Wiley, New York, 2000)
17. M.L. Culpepper, G. Anderson, Design of a low-cost nano-manipulator which utilizes a monolithic, spatial compliant mechanism. Precis. Eng. **28**(4), 469–482 (2004)
18. T. Dahlberg, Procedure to calculate deflections of curved beams. Int. J. Eng. Educ. **20**(3), 503–513 (2004)
19. N.B. Dahotre, S.P. Harimkar, *Laser Fabrication and Machining of Materials* (Springer, New York, 2008)
20. S. Devasia, E. Eleftheriou, S.O.R. Moheimani, A survey of control issues in nanopositioning. IEEE Trans. Control Syst. Technol. **15**, 802–823 (2007)
21. G.K. Fedder, Simulation of microelectromechanical systems. Ph.D. thesis, Electrical Engineering and Computer Sciences, University of California at Berkeley (1994)
22. A.J. Fleming, A megahertz bandwidth dual-amplifier for driving piezoelectric actuators and other highly capacitive loads. Rev. Sci. Instrum. **80**(1–7), 104701 (2009)
23. A. Fleming, Dual-stage vertical feedback for high-speed scanning probe microscopy. IEEE Trans. Control Syst. Technol. **19**(1), 156–165 (2011)
24. A.J. Fleming, K.K. Leang, *Design, Modeling and Control of Nanopositioning Systems* (Springer, London, 2014).
25. P. Gao, S. Swei, Z. Yuan, A new piezodriven precision micropositioning stage utilizing flexure hinges. Nanotechnology **10**, 394–398 (1999)

26. D.C. Handley, T.F. Lu, Y.K. Yong, Workspace investigation of a 3dof compliant micro-motion stage, in *The Eighth International Conference on Control, Automation, Robotics and Vision, ICARCV*, pp. 1279–1284 (2004)
27. K.H. Ho, S.T. Newman, S. Rahimifard, R.D. Allen, State of the art in wire electrical discharge machining (WEDM). Int. J. Mach. Tools Manuf. **44**(12–13), 1247–1259 (2004)
28. L. Howell, *Compliant Mechanisms*. (Wiley, New York, 2001)
29. A.D.L. Humphris, B. Zhao, D. Catto, J.P. Howard-Knight, P. Kohli, J.K. Hobbs, High speed nano-metrology. Rev. Sci. Instrum. **82**(4), 043710–5 (2011)
30. D.J. Inman, *Engineering Vibration*. Prentice Hall Inc., Upper Saddle River (1996)
31. M.S. Johannes, D.G. Cole, R.L. Clark, Atomic force microscope based nanofabrication of master pattern molds for use in soft lithography. Appl. Phys. Lett. **91**(12), 123111 (2007)
32. M. Jouaneh, R. Yang, Modeling of flexure-hinge type lever mechanisms. Precis. Eng. **27**, 407–418 (2003)
33. B.J. Kenton, Design, characterization, and control of a high-bandwidth serial-kinematic nanopositioning stage for scanning probe microscopy applications. Master's thesis, University of Nevada, Reno (2010)
34. B.J. Kenton, A.J. Fleming, K.K. Leang, Compact ultra-fast vertical nanopositioner for improving scanning probe microscope scan speed. Rev. Sci. Instrum. **82**(12), 123703 (2011)
35. B.J. Kenton, K.K. Leang, Design and control of a three-axis serial-kinematic high-bandwidth nanopositioner. IEEE/ASME Trans. Mechatron. **17**(2), 356–368 (2012)
36. D. Kim, D. Kang, J. Shim, I. Song, D. Gweon, Optimal design of a flexure hinge-based XYZ atomic force microscopy scanner for minimizing abbe errors. Rev. Sci. Instrum. **76**, 073706 (2005)
37. J.H. Kindt, G.E. Fantner, J.A. Cutroni, P.K. Hansma, Rigid design of fast scanning probe microscopes using finite element analysis. Ultramicroscopy **100**, 259–265 (2004)
38. K.K. Leang, A.J. Fleming, High-speed serial-kinematic AFM scanner: design and drive considerations. Asian J. Control (Special issue on Advanced Control Methods for Scanning Probe Microscopy Research and Techniques) **11**(2), 144–153 (2009)
39. D.J. Leo, *Engineering Analysis of Smart Material Systems*. (Wiley, Hoboken, 2007)
40. D. Leonard, M. Krishnamurthy, C.M. Reaves, S.P. Denbaars, P.M. Petroff, Direct formation of quantum-sized dots from uniform coherent islands of InGaAs on GaAs surfaces. Appl. Phys. Lett. **63**(23), 3203–3205 (1993)
41. Y. Li, Q. Xu, Modeling and performance evaluation of a flexure-based xy parallel micromanipulator. Mech. Mach. Theory **44**(12), 2127–2152 (2009)
42. Y. Li, Q. Xu, Development and assessment of a novel decoupled xy parallel micropositioning platform. IEEE/ASME Trans. Mechatron. **15**(1), 125–135 (2010)
43. H.C. Liaw, B. Shirinzadeh, Neural network motion tracking control of piezo-actuated flexure-based mechanisms for micro-/nanomanipulation. IEEE/ASME Trans. Mechatron. **14**(5), 517–527 (2009). doi:10.1109/TMECH.2009.2005491
44. N. Lobontiu, *Compliant Mechanisms: Design of Flexure Hinges* (CRC Press, Boca Raton, 2003)
45. N. Lobontiu, E. Garcia, Analytical model of displacement amplification and stiffness optimization for a class of flexure-based compliant mechanisms. Comput. Struct. **81**, 2797–2810 (2003)
46. N. Lobontiu, E. Garcia, Two-axis flexure hinges with axially-collocated and symmetric notches. Comput. Struct. **81**, 1329–1341 (2003)
47. N. Lobontiu, J.S.N. Paine, E. Garcia, M. Goldfarb, Corner-filleted flexure hinges. Trans. ASME J. Mech. Des. **123**, 346–352 (2001)
48. N. Lobontiu, J.S.N. Paine, E. O'Malley, M. Samuelson, Parabolic and hyperbolic flexure hinges: flexibility, motion precision and stress characterization based on compliance closed-form equations. Precis. Eng. **26**, 183–192 (2002)
49. T.F. Lu, D.C. Handley, Y.K. Yong, C. Eales, A three-DOF compliant micromotion stage with flexure hinges. Ind. Robot **31**(4), 355–361 (2004)

50. H.W. Ma, Y. Shao-Ming, L.Q. Wang, Z. Zhong, Analysis of the displacement amplification ratio of bridge-type flexure hinge. Sens. Actuators A Phys. **132**(2), 730–736 (2006)
51. I.A. Mahmood, S.O.R. Moheimani, Fast spiral-scan atomic force microscopy. Nanotechnology **20**(36), 365503 (2009)
52. I.A. Mahmood, S.O.R. Moheimani, B. Bhikkaji, A new scanning method for fast atomic force microscopy. IEEE Trans. Nanotechnol. **10**(2), 203–216 (2010)
53. M. Maroufi, A.G. Fowler, A. Bazaei, S.O.R. Moheimani, High-stroke silicon-on-insulator MEMS nanopositioner: control design for non-raster scan AFM. Rev. Sci. Instrum. **86**(2), 023705 (12pp) (2015)
54. R. Merry, N. de Kleijn, M. van de Molengraft, M. Steinbuch, Using a walking piezo actuator to drive and control a high-precision stage. IEEE/ASME Trans. Mechatron. **14**(1), 21–31 (2009)
55. S. Miyake, M. Wang, J. Kim, Silicon nanofabrication by atomic force microscopy-based mechanical processing. J. Nanotechnol. **2014**, 102404 (2014)
56. A. Mohammadi, A.G. Fowler, Y.K. Yong, S.O.R. Moheimani, A feedback controlled mems nanopositioner for on-chip high-speed afm. IEEE J. Microelectromech. Syst. **23**(3), 610–619 (2014)
57. S.O.R. Moheimani, Invited review article: accurate and fast nanopositioning with piezoelectric tube scanners: Emerging trends and future challenges. Rev. Sci. Instrum. **79**(7), 071101 (2008)
58. B. Mokaberi, A.A.G. Requicha, Compensation of scanner creep and hysteresis for AFM nanomanipulation. IEEE Trans. Autom. Sci. Eng. **5**(2), 197–206 (2008)
59. Newcastle Innovation Ltd., Industry Development Centre, University Drive, Callaghan NSW 2308, Australia: PiezoDrive PDL200 Manual, www.piezodrive.com
60. S.R. Park, A mathematical approach for analyzing ultra precision positioning system with compliant mechanism. J. Mater. Res. Process. Technol. **164–165**, 1584–1589 (2005)
61. J.M. Paros, L. Weisbord, How to design flexure hinges. Mach. Des. **37**, 151–156 (1965)
62. J. Peng, X. Chen, Modeling of piezoelectric-driven stick-slip actuators. IEEE/ASME Trans. Mechatron. **16**(2), 394–399 (2011)
63. A.P. Pisano, Y.H. Cho, Mechanical design issues in laterally-driven microstructures. Sens. Actuators A Phys. **23**(1–3), 1060–1064 (1990). Proceedings of the 5th International Conference on Solid-State Sensors and Actuators and Eurosensors {III}
64. A.B. Puri, B. Bhattacharyya, An analysis and optimisation of the geometrical inaccuracy due to wire lag phenomenon. Int. J. Mach. Tools Manuf. **43**, 151–159 (2003)
65. M. Rakotondrabe, Y. Haddab, P. Lutz, Development, modeling, and control of a micro-/nanopositioning 2-dof stick-slip device. IEEE/ASME Trans. Mechatron. **14**(6), 733–745 (2009)
66. J.W. Ryu, D.G. Gweon, Error analysis of a flexure hinge mechanism induced by machining imperfection. Precis. Eng. **21**, 83–89 (1997)
67. S. Salapaka, A. Sebastian, J.P. Cleveland, M.V. Salapaka, High bandwidth nano-positioner: A robust control approach. Rev. Sci. Instrum. **73**(9), 3232–3241 (2002)
68. R.B. Salazar, A. Shovsky, H. Schonherr, G.J. Vancso, Dip-pen nanolithography on (bio)reactive monolayer and block-copolymer platforms: Deposition of lines of single macromolecules. Small **2**(11), 1274–1282 (2006)
69. G. Schitter, K.J. Åstrom, B. DeMartini, P.J. Thurner, K.L. Turner, P.K. Hansma, Design and modeling of a high-speed AFM-scanner. IEEE Trans. Control Syst. Technol. **15**(5), 906–915 (2007)
70. G. Schitter, P.J. Thurner, P.K. Hansma, Design and input-shaping control of a novel scanner for high-speed atomic force microscopy. Mechatronics **18**, 282–288 (2008)
71. F. Scire, E. Teague, Piezodriven 50-μm range stage with subnanometer resolution. Rev. Sci. Instrum. **49**(12), 1735–1740 (1978)
72. See Asylum Research: http://asylumresearch.com for information on flexure-based nanopositioning platforms

73. See Mad City Labs Inc: http://www.madcitylabs.com/ for information on flexure-based nanopositioning applications
74. See Park Systems: www.parkafm.com for information on the use of flexure-based nanopositioning platforms in commercially available AFMs
75. E. Shamoto, T. Moriwaki, Ultraprecision diamond cutting of hardened steel by applying elliptical vibration cutting. CIRP Ann. Manuf. Technol. **48**(1), 441–444 (1999)
76. M. Shiga, Invar alloys. Curr. Opin. Solid State Mater. Sci. **1**(3), 340–348 (1996). doi:10.1016/S1359-0286(96)80023-4
77. V. Shilpiekandula, K. Youcef-Toumi, Dynamic modeling and performance trade-offs in flexure-based positioning and alignment systems, Chapter 25, in *Motion Control*, ed. by F. Casolo (InTech, Croatia, 2010), pp. 481–496
78. S.T. Smith, *Flexures: Elements of Elastic Mechanisms*. (Gordon and Breach, New York, 2000)
79. G.T. Smith, *Cutting Tool Technology: Industrial Handbook*. (Springer, London, 2008)
80. S.T. Smith, V.G. Badami, J.S. Dale, Y. Xu, Elliptical flexure hinges. Rev. Sci. Instrum. **68**(3), 1474–1483 (1997)
81. T.A. Spedding, Z.Q. Wang, Study on modeling of wire EDM process. J. Mater. Process. Technol. **69**, 18–28 (1997)
82. D.J. Taatjes, A.S. Quinn, J.H. Rand, B.P. Jena, Atomic force microscopy: high resolution dynamic imaging of cellular and molecular structure in health and disease. J. Cell. Physiol. **228**(10), 1949–1955 (2013)
83. The World of Micro- and NanoPositioning (PI catalog). Physik Instrumente, Karlsruhe (2005)
84. S.P. Timoshenko, *History of Strength of Materials*. (McGraw-Hill Book Company, New York, 1953)
85. A.A. Tseng, Advancements and challenges in development of atomic force microscopy for nanofabrication. Nano Today **6**(5), 493–509 (2011)
86. A.A. Tseng, S. Jou, A. Notargiacomo, T. Chen, Recent developments in tip-based nanofabrication and its roadmap. J. Nanosci. Nanotechnol. **8**, 2167–2186 (2008)
87. Y.M. Tseytlin, Notch flexure hinges: an effective theory. Rev. Sci. Instrum. **73**(9), 3363–3368 (2002)
88. T. Tuma, J. Lygeros, V. Kartik, A. Sebastian, A. Pantazi, High-speed multiresolution scanning probe microscopy based on Lissajous scan trajectories. Nanotechnology **23**(18), 185501 (2012)
89. S. Wadikhaye, Y.K. Yong, S.O.R. Moheimani, B. Bhikkaji, Control of a piezoelectrically actuated high-speed serial kinematic AFM nanopositioner. Smart Mater. Struct. **22**(2), 025030 (12pp.) (2014)
90. C. Wang, V.K. Yadavalli, Investigating biomolecular recognition at the cell surface using atomic force microscopy. Micron **60**(0), 5–17 (2014)
91. D.L. White, O.R. Wood, Novel alignment system for imprint lithography. Rev. Sci. Instrum. **18**(6), 3552–3556 (2000)
92. K. Wiesauer, G. Springholz, Fabrication of semiconductor nanostructures by nanoindentation of photoresist layers using atomic force microscopy. J. Appl. Phys. **88**(12), 7289–7297 (2000)
93. P. Xu, Y. Jingjun, Z. Guanghua, B. Shusheng, Y. Zhiwei, Analysis of rotational precision for an isosceles-trapezoidal flexural pivot. J. Mech. Des. **130**(5), 052302 (9pp.) (2008). doi:10.1115/1.2885507
94. Y.K. Yong, A. Fleming, Piezoelectric actuators with integrated high-voltage power electronics. IEEE/ASME Trans. Mechatron. **20**(2), 611–617 (2015)
95. Y.K. Yong, T.F. Lu, The effect of the accuracies of flexure hinge equations on the output compliances of planar micro-motion stages. Mech. Mach. Theory **43**, 347–363 (2008)
96. Y.K. Yong, T.F. Lu, Kinetostatic modeling of 3-RRR compliant micro-motion stages with flexure hinges. Mech. Mach. Theory **44**(6), 1156–1175 (2009)
97. Y.K. Yong, S.O.R. Moheimani, A compact XYZ scanner for fast atomic force microscopy in constant force contact mode, in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, Montreal (2010)

98. Y.K. Yong, S.O.R. Moheimani, Design of an inertially counterbalanced Z-nanopositioner for high-speed atomic force microscopy. IEEE/ASME Trans. Nanotechnol. **12**(2), 137–145 (2013)
99. Y.K. Yong, T.F. Lu, D.C. Handley, Review of circular flexure hinge design equations and derivation of empirical formulations. Precis. Eng. **32**(2), 63–70 (2008)
100. Y.K. Yong, S. Aphale, S.O.R. Moheimani, Design, identification and control of a flexure-based XY stage for fast nanoscale positioning. IEEE Trans. Nanotechnol. **8**(1), 46–54 (2009)
101. Y.K. Yong, K. Liu, S.O.R. Moheimani, Reducing cross-coupling in a compliant XY nanopositioning stage for fast and accurate raster scanning. IEEE Trans. Control Syst. Technol. **18**(5), 1172–1179 (2010)
102. Y.K. Yong, S.O.R. Moheimani, I.R. Petersen, High-speed cycloid-scan atomic force microscopy. Nanotechnology **21**(36), 365503 (2010)
103. Y.K. Yong, B. Bhikkaji, S.O.R. Moheimani, Analog control of a high-speed atomic force microscope scanner, in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, Budapest (2011)
104. Y.K. Yong, S.O.R. Moheimani, B.J. Kenton, K.K. Leang, Invited review article: high-speed flexure-guided nanopositioning: mechanical design and control issues. Rev. Sci. Instrum. **83**(12), 121101 (2012)
105. Y.K. Yong, B. Bhikkaji, S.O.R. Moheimani, Design, modeling and FPAA-based control of a high-speed atomic force microscope nanopositioner. IEEE/ASME Trans. Mechatron. **18**(3), 1060–1071 (2013)
106. Y.K. Yong, A. Bazaei, S.O.R. Moheimani, Video-rate Lissajous-scan atomic force microscopy. IEEE Trans. Nanotechnol. **13**(1), 85–93 (2014)
107. W.C. Young, R.G. Budynas, *Roark's Formula for Stress and Strain*, 7th edn. (McGraw-Hill, New York, 2002)

Chapter 4
Parallel-Kinematic Nanopositioning Stages Based on Roberts Mechanism

Sicong Wan and Qingsong Xu

Abstract This chapter presents the design of parallel-kinematic nanopositioning stages with large workspace and low crosstalk. Based on compliant Roberts mechanism, the design of multi-axis parallel compliant stage is synthesized. A new XY nanopositioning stage is studied in detail. Pseudo-rigid-body model (PRBM) is developed to build the quantitative models of the compliant Roberts mechanisms. In addition, finite-element analysis (FEA) is carried out to validate its performance. A prototype is fabricated and tested through experimental studies. Results show that the XY stage delivers a work range larger than 12 mm in each axis. Moreover, the parasitic motion of the stage in the non-working direction is less than 1.7 % of the motion stroke. It indicates that the designed stage meets the requirements of a large stroke and high precision linear guiding mechanism, which demonstrates the feasibility of the proposed design ideas using the Roberts mechanism. Moreover, a feedback control using PID control algorithm is implemented to demonstrate the positioning performance of the developed XY stage. The reported ideas can also be extended to the design and control of other types of micro-/nanopositioning systems.

4.1 Introduction

With the development of technology, the micropositioning and displacement technology has played an important role in the defense industry, microelectronics engineering, aerospace technology and biological engineering, and other fields [9, 12, 22, 28, 29, 32, 33]. It directly affects the whole machinery industry technology and technique development, determines the quality of machining accuracy and performance of products, and also affects the reliability and stability of the device. Compliant mechanism can transmit motion and force by using the material's elastic deformation capability. As compared with rigid-body mechanism, compliant mechanism eliminates the problems such as clearance, friction, and wears [8]. In the last few years, many scholars have proposed different flexible displacement

S. Wan • Q. Xu (✉)
Department of Electromechanical Engineering, Faculty of Science and Technology, University of Macau, Avenida da Universidade, Taipa, Macau, China
e-mail: mb25541@umac.mo; qsxu@umac.mo

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_4

mechanism. Generally, the compliant mechanism whose workspace is larger than 1 mm is known as a large-stroke compliant mechanism, while the compliant mechanism whose workspace is less than 1 mm is called a small-stroke compliant mechanism. According to the flexible hinge type, the notch-type flexible hinge mechanism is difficult to obtain a large-stroke movement because of the limited flexible deformation. However, according to the flexible structure types, stiffness value of the flexible mechanism which consists of redundant symmetric structures is very sensitive to the driving force. Therefore, it is also difficult to achieve a large-stroke motion. Appropriately increasing the size of the flexible structure can increase the workspace, but the installation space will limit the size of the stage.

On the other hand, because it is hard to make a precision micropositioning system with both a large workspace and multi-degree of freedom at the same time, the applications of compliant mechanism are restricted [13, 21, 27, 30]. For instance, many applications require a compliant mechanism which has a larger range of work. To name a few, the micro-coordinate measuring machines, adaptive adjustment mechanism in astronomical telescope, MEMS processing systems, binary optical processing equipment, and so on [6, 16]. Hence, the large-stroke and multi-DOF micropositioning stage has drawn extensive attentions of some recent research. For example, Tan et al. [18] proposed a multi-leaf configuration of expanded-motion-range flexure hinges. Zhao et al. [34] designed a new large-deflection pivot based on curved flexure element. Xu [25] designed a novel compact long-stroke precision positioning stage based on flexible leaf springs. Trease [20] proposed a large-displacement compliant joint, and Choi et al. [5] designed a large-displacement precision XY positioning stage by using the cross strip flexure joints and over-constrained mechanism.

According to previous work [2, 7, 15, 17, 19, 26], the motion accuracy of such mechanisms is generally about 1 % of its workspace. In traditional design, there is no strict compensation for the coupled displacement among axes. So, such structures exhibit a large amount of coupling displacement. For example, Lin and Lee [15] put forward a comprehensive flexible structure of a large-stroke compliant mechanism, which is a flexible mechanism based on pseudo-symmetrical structure. Simulation results show that its workspace is 2.34 mm. However, the amount of coupling displacement in non-working direction is almost 390 μm. That is, the cross-axis coupling amounts to 16.6 %. Therefore, to achieve a precise positioning, it is necessary to reduce the crosstalk value.

What's more, it is challenging to devise a compliant micropositioning stage with a stroke longer than 10 mm in each working axis. One reason lies in that the size of the flexible component is the decisive factor which determines its workspace directly [24]. It is not easy to devise an XY stage have both large workspace and compact dimension simultaneously [24]. In recent research, a compact stage is designed in [11], which has a dimension of $100 \times 100\,\text{mm}^2$. Yet, the stage can just provide a workspace of $2 \times 2\,\text{mm}^2$. In contrast, the XY stage designed in [14] produces a workspace of $20 \times 20\,\text{mm}^2$, but it is contributed by a large physical dimension of $432 \times 432\,\text{mm}^2$.

To this end, the conceptual design of a new decoupled XY micropositioning stage is proposed in this research. It can deliver a large stroke and possesses a relatively compact structure thanks to a two-layer design. Specifically, the Roberts mechanism is introduced to design a linear guiding mechanism. Considering that a single Roberts mechanism can only provide the linear motion of a single point, two Roberts mechanisms working in parallel is employed to construct a flexible linear guiding mechanism. By cascading two newly proposed linear compliant guiding mechanisms, a symmetrical structure design yields a novel large-stroke compliant parallel-kinematic XY micropositioning stage based on Roberts mechanism. To the knowledge of the authors, no work has been conducted to devise a multi-axis micropositioning stage using the Roberts mechanism in the literature.

The main contribution of this work is the conceptual design and verification of a large-stroke parallel-kinematic compliant XY micropositioning stage based on Roberts mechanism, which has not been attempted in previous literature. The remaining parts of the paper are organized as follows. Section 4.2 presents the conceptual design of the new XY stage. Analytical modeling of the XY stage is conducted in Sect. 4.3. These models are verified by conducting finite-element analysis (FEA) simulations in Sect. 4.4. Section 4.5 describes a fabricated prototype and experimental investigations. Concluding remarks are summarized in Sect. 4.6.

4.2 Conceptual Design of Large-Range XY Compliant Mechanism

In this section, the traditional Roberts mechanism is introduced. As a traditional rigid-body mechanism, it can convert rotational motion into straight-line motion approximately. Based on this mechanism, two novel compliant parallel linear guiding mechanisms are proposed.

4.2.1 Background of Roberts Mechanism

Roberts mechanism was first proposed by Richard Roberts (1789–1864) [1]. Its schematic is shown in Fig. 4.1a. When the link lengths are designed as R1 = R3 and the central blue triangle is isosceles (i.e., the two non R2 sides are of equal length), the trajectory of point P will become an approximate straight line, as shown in Fig. 4.1b.

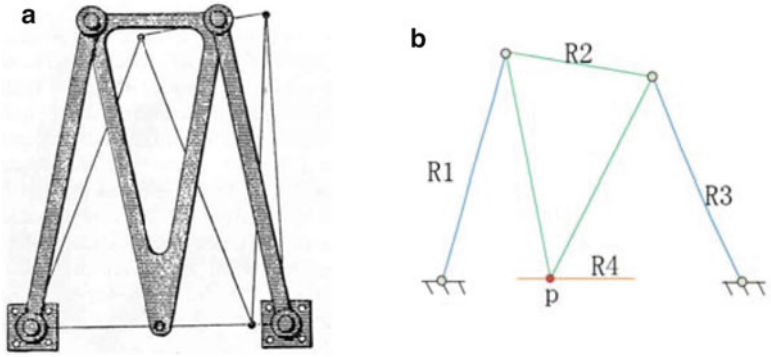


Fig. 4.1 Schematic diagram of the Roberts mechanism. (**a**) Traditional Roberts mechanism [23]; (**b**) structure of the Roberts mechanism

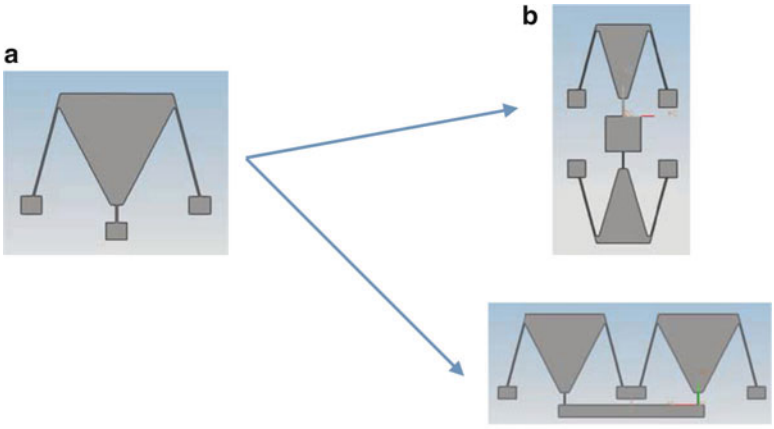


Fig. 4.2 (**a**) Flexure Roberts mechanism; (**b**) two new compliant Roberts linear guiding mechanisms

4.2.2 Proposal of New Compliant Linear Guiding Mechanism

By replacing the conventional kinematic pairs of traditional rigid-body Roberts mechanism with flexible leaf springs, a new kind of compliant linear guiding mechanism is obtained, as shown in Fig. 4.2a. However, a single Roberts mechanism can only provide a single point of linear motion. To overcome this issue, two Roberts mechanisms connected in parallel can be employed to generate a compliant parallel linear guiding mechanism. According to the joining ways, two different mechanisms are obtained as shown in Fig. 4.2b.

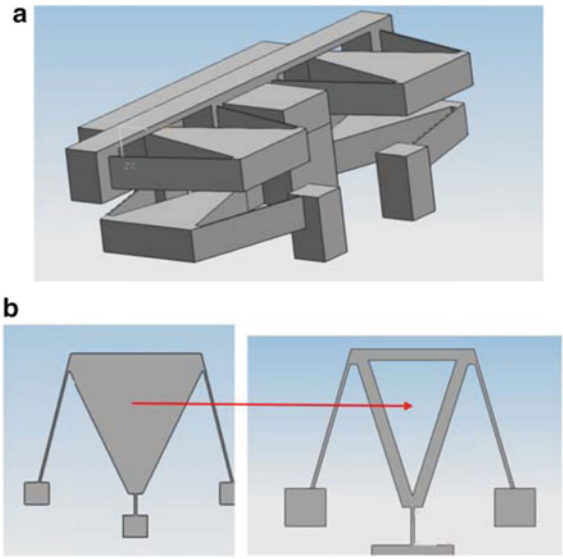


Fig. 4.3 A novel serial XY flexure mechanism. (**a**) XY mechanism; (**b**) lightweight optimization process

4.2.3 Design of New Compliant Serial XY Linear Guiding Mechanism

Based on the two new flexible linear guiding mechanisms, a novel XY mechanism is designed as shown in Fig. 4.3a. It is observed that the two compliant Roberts mechanisms are connected in series. In addition, the proposed design is a two-layer structure, and it can yield decoupling motion in X axis and Y axis. The two-layer design enables the generation of a relatively compact dimension design.

In order to reduce the weight and to improve the resonant frequency, a lightweight optimization process is applied to the serial XY guiding mechanism as illustrated in Fig. 4.3b. The resulted serial XY mechanism is employed in the following design of an XY stage.

4.2.4 Design of New Compliant Parallel-Kinematic XY Micropositioning Stage

As a conceptual design, four serial XY linear guiding mechanisms are used to construct a parallel XY micropositioning stage, which owns a 4-PP (P stands for prismatic joint) parallel mechanism. The CAD model of the overall structure for an XY stage is shown in Fig. 4.4. This structure meets the requirement of the integrated machining and easy assembly. In addition, the lightweight optimization process reduces the mass of the platform while improves the dynamic performance of the

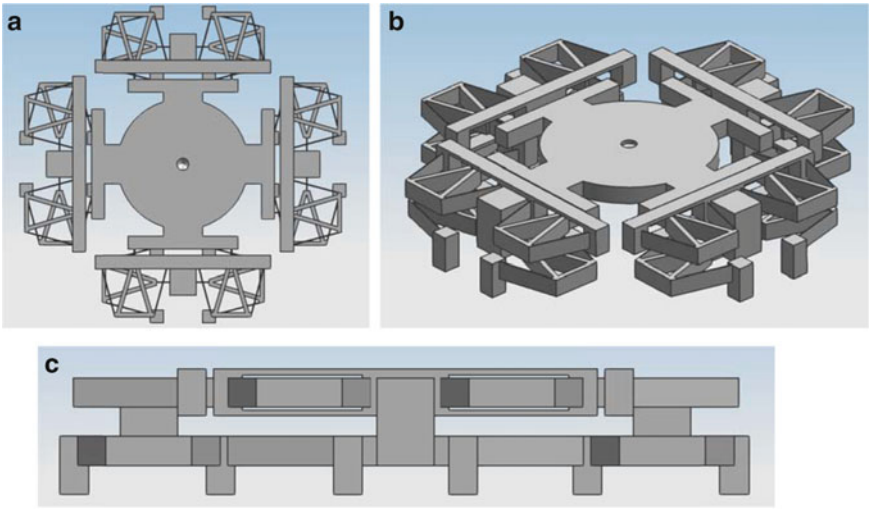


Fig. 4.4 A new compliant parallel XY large-stroke micropositioning stage. (**a**) Top view; (**b**) isotropic view; (**c**) side view

system. Moreover, the idea of symmetrical structure is also adopted in the stage design. Such a symmetrical design can compensate for the coupling displacement to some extent, as revealed in the later discussion.

The specific dimension design is introduced in the following section by establishing theoretical models.

4.3 Analytical Modeling of the XY Stage

4.3.1 PRBM of the Compliant Roberts Mechanism

As discussed above, the parallel-kinematic XY compliant stage consists of four serial XY compliant Roberts mechanisms. For each single compliant Roberts mechanism, the kinematic principle is illustrated in Fig. 4.5. When an external force is applied to the coupler point in the extension direction, the two flexure beams are deformed to generate an approximate straight-line motion at the coupler point.

To calculate the stiffness of a Roberts mechanism, the pseudo-rigid-body model (PRBM) can be used. The concept of PRBM is to model the force-displacement relationship of a compliant element using an equivalent rigid-body mechanism with the compliance modeled as torsional springs around the joints. This allows the vast knowledge of traditional mechanism kinematics to be applied to the design of a compliant element.

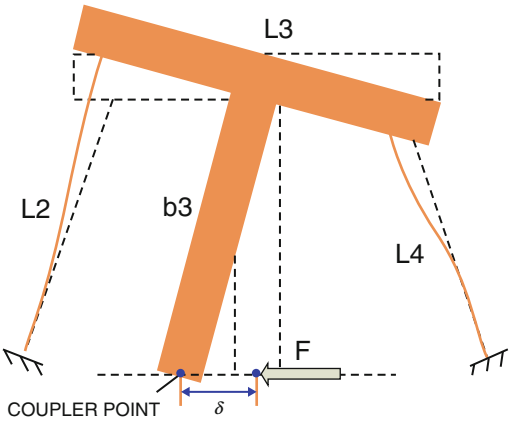


Fig. 4.5 Kinematic principle of a flexible Roberts mechanism

The PRBM of a single beam is shown in Fig. 4.6a. In addition, the PRBM of a single flexure Roberts mechanism is shown in Fig. 4.6b, where the two flexure beams are replaced by four identical torsional springs and two rigid links [8]. The overall stiffness for this group of flexure joints can be obtained through

$$K = \frac{F}{\Delta x}, \tag{4.1}$$

where Δx and F are the displacement and applied force for the movable plate, respectively. These two parameters are obtained by the following equations:

$$\Delta x = r_2(\cos\theta_2 - \cos\theta_{20}) + \frac{r_3}{2}(\cos\theta_3 - 1) + b_3\sin\theta_3 \tag{4.2}$$

$$F = \frac{4K_\theta[(2-h_{32})\Delta\theta_2 + (2h_{42}-h_{32})\Delta\theta_4 - (1+h_{42}-h_{32})\Delta\theta_3]}{r_2\sin\theta_2 + h_{32}(\frac{r_3}{2}\sin\theta_3 - b_3\cos\theta_3)}, \tag{4.3}$$

where b_3 is the perpendicular distance from the center of link 3 to the coupler point (see Fig. 4.6b). r_i, θ_i, and θ_{i0} $(i = 2, 3, 4)$ are the pseudo-rigid-body (PRB) length, final angle, and initial angle of link i, respectively. The values for r_2 and r_3 are defined as γl, as shown in Fig. 4.6b. In addition, h_{32}, h_{42}, and K_θ are the kinematic coefficients defined by the following equations [3]:

$$h_{32} = \frac{\delta\theta_3}{\delta\theta_2} = \frac{\omega_3}{\omega_2} = \frac{r_2\sin(\theta_4-\theta_2)}{r_3\sin(\theta_3-\theta_4)} \tag{4.4}$$

$$h_{42} = \frac{\delta\theta_3}{\delta\theta_2} = \frac{\omega_4}{\omega_2} = \frac{r_2\sin(\theta_3-\theta_2)}{r_4\sin(\theta_3-\theta_4)} \tag{4.5}$$

$$K_\theta = \gamma K_\Theta \frac{Ech^3}{6l}, \tag{4.6}$$

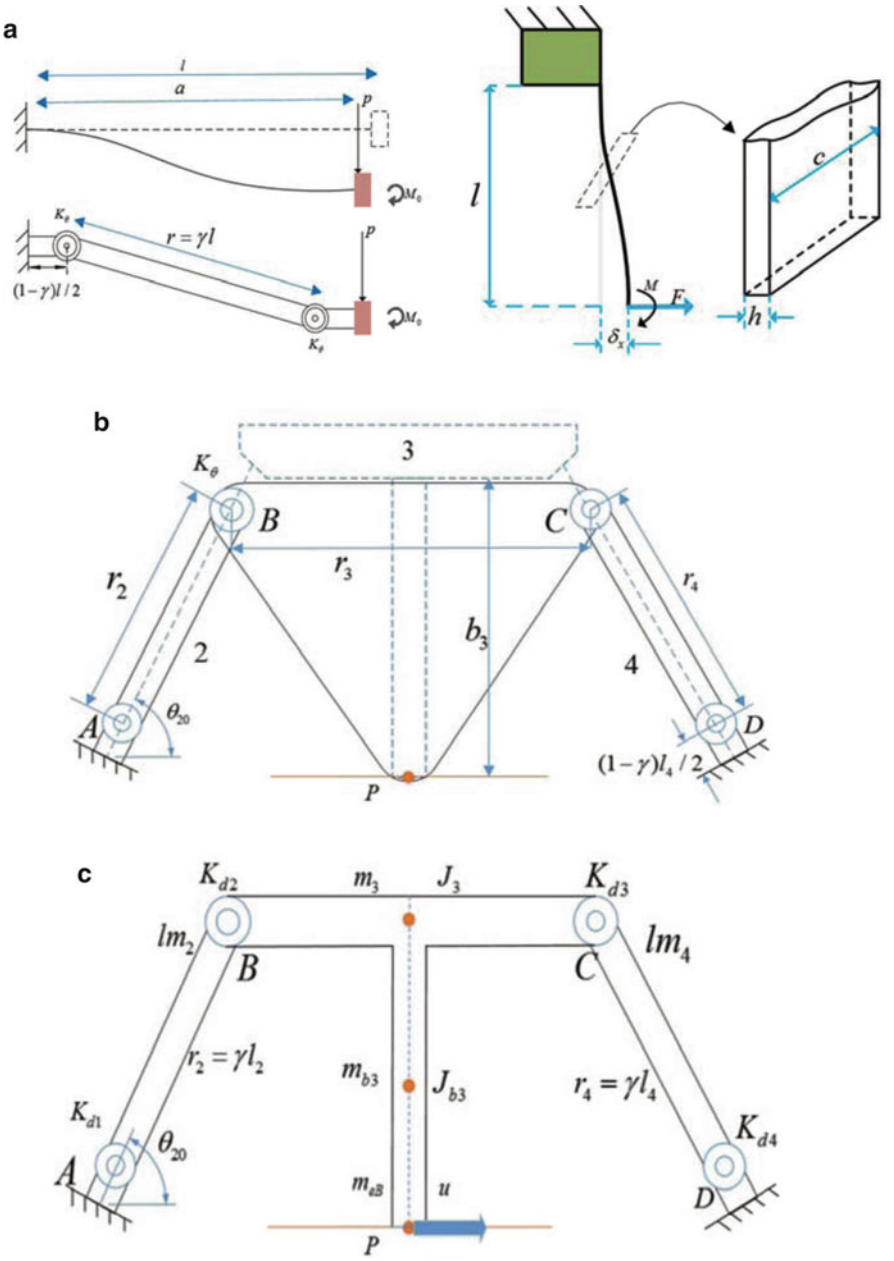


Fig. 4.6 The PRBM of compliant Roberts mechanism. (**a**) The PRBM of a single beam; (**b**) the PRBM of a single Roberts mechanism; (**c**) RBM for the Roberts mechanism

where γ is the characteristic radius factor and K_Θ is the stiffness coefficient. For this PRBM approximation, the values of γ and K_Θ are set as 0.852 and 2.65, respectively [10]. The terms $\Delta\theta_2$, $\Delta\theta_3$, and $\Delta\theta_4$ are the differences of link angles with respect to their initial positions.

In order to determine the required force for a given displacement, the instantaneous angles of the three links should be known. Given an input displacement, all the angles can be found by solving (4.2) and the explicit four-bar equations simultaneously [8]. Then, these angles can be substituted into (4.3) to find the required force. The resulting stiffness K of the flexure joint is determined by (4.1).

For an XY stage design as shown in Fig. 4.4a, a stiffness model is derived as shown in Fig. 4.7. Then, the stiffness of the XY stage seen at the input end can be computed as

$$K_{\text{in1}} = 8K. \tag{4.7}$$

So, $F_{\max}$ in one-sided direction can be calculated as follow.

$$F_{\max} \geq 8K \times D_{\max}. \tag{4.8}$$

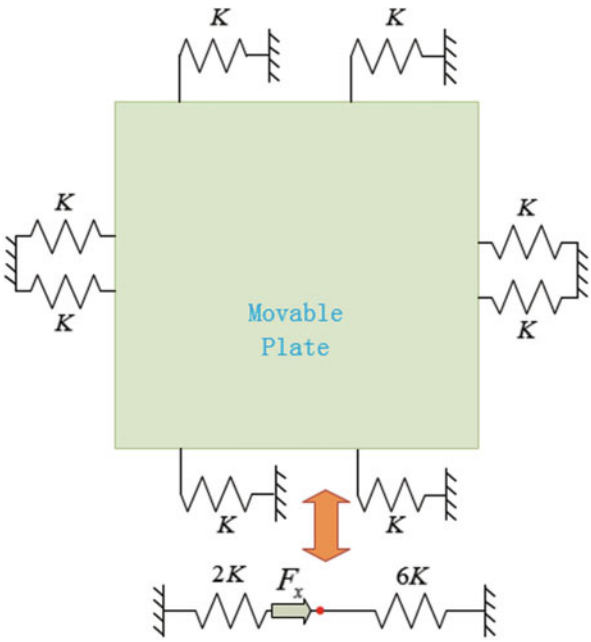


Fig. 4.7 Stiffness model of the XY stage

4.3.2 Dynamics Modeling

For a single compliant Roberts mechanism, four torsional springs ($K_{d1} = K_{d2} = K_{d3} = K_{d4}$) and two lumped masses of the two flexible links ($lm_2 = lm_4$) are added in the PRBM, as shown in Fig. 4.6c [1, 2, 4, 8]. Considering that $m_2 = m_4$, $J_2 = J_4$, and $r_2 = r_4$, lm_2 can calculate as:

$$J_i = \frac{1}{12} m_i l_i \tag{4.9}$$

$$lm_2 = lm_4 = \frac{m_2}{4} + \frac{J_2}{r_2^2}, \tag{4.10}$$

where $m_2 = m_4 = l \times c \times h \times \rho_{\mathrm{AL}-7075}$ with $\rho_{\mathrm{AL}-7075}$ denoting the density of the material.

The kinetic energy for the PRBM system can be expressed as follows [4]:

$$\begin{aligned} T &= \frac{1}{2} m_{eB} u^2 \\ &= \frac{1}{2} lm_2 v_B^2 + \frac{1}{2} lm_4 v_C^2 + \left(\frac{1}{2} m_3 v_{m3}^2 + \frac{1}{2} J_3 \omega_3^2 \right) + \left(\frac{1}{2} m_{b3} v_{b3}^2 + \frac{1}{2} J_{b3} \omega_3^2 \right), \end{aligned} \tag{4.11}$$

where m_{eB} is the equivalent mass of a single compliant Roberts mechanism, m_i and J_i are the mass and moment of inertia of link i ($i = 3, b$), respectively. lm_2 and lm_4 are the lumped masses of the two flexible links, which are given in [31]. u is the velocity of the coupler point, ω_3 is the angular velocity of link 3, and v_B, v_C, v_{m3}, and v_{b3} are the velocities of the corresponding points, respectively.

From the kinematic analysis of the rigid-body motion of a four-bar linkage, it is known that

$$V_B = V_C = f(\theta_2, \dot{u}) \tag{4.12}$$

$$V_{m3} = g(\theta_2, \dot{u}) \tag{4.13}$$

$$V_{b3} = h(\theta_2, \dot{u}) \tag{4.14}$$

$$\omega_3 = q(\theta_2, \dot{u}). \tag{4.15}$$

Therefore, the equivalent mass of a single compliant Roberts mechanism can be obtained as [1]:

$$m_{eB} = \alpha(lm_2 + lm_4) + \beta m_3 + \psi m_{b3} + \frac{1}{\chi^2}(J_3 + J_{b3}), \tag{4.16}$$

where

$$\chi = \frac{r_3}{2}\tan\theta_2 + b_3 \tag{4.17}$$

$$\alpha = \frac{r_3^2}{4\chi^2\cos^2\theta_2} \tag{4.18}$$

$$\beta = \left(1 - \frac{b_3}{\chi}\right)^2 \tag{4.19}$$

$$\psi = \left(1 - \frac{b_3}{2\chi}\right)^2. \tag{4.20}$$

In order to generate a high bandwidth of the control system, a high resonant frequency of the stage is required. Based on Lagrange's equation, the free motion of the XY stage can be described by the dynamic equation

$$(8m_{eB} + m_{\text{mov}}) \times \ddot{q}(t) + K_{\text{in}}\, q(t) = 0, \tag{4.21}$$

where m_{mov} is the mass of the movable plate of the XY stage and $q = [x\ y]^T$ is the displacement of the movable plate in two working directions.

Based on the theory of vibration, the modal equation can be derived as

$$[K_{\text{in}} - \omega_i^2(8m_{eB} + m_{\text{mov}})]\Phi_i = 0, \tag{4.22}$$

where the eigenvector Φ_i (for $i = 1$ and 2) represents a mode shape and the eigenvalue describes the corresponding natural cyclic frequency, which can be obtained by solving the characteristic equation. Because the mass matrix and stiffness matrix of the stage are positive definite, the natural frequency of this XY stage in each working axis is computed as

$$f_i = \frac{1}{2\pi}\sqrt{\frac{K_{\text{in}}}{8m_{eB} + m_{\text{mov}}}}. \tag{4.23}$$

4.4 Model Verification with FEA Simulations

In this work, the XY stage is designed to generate a stroke over ± 7 mm in each working direction as an illustration. For the convenience of calculation, the angle parameters are selected as $\theta_{20} = \theta_{40} = 75°$, $r_3 = 30$ mm, and $\theta_{30} = 180°$. Moreover, the natural frequency of the system is constrained to be greater than 40 Hz. Such constraint is imposed by using the dynamic model (4.16). Based on the design criteria (4.2), (4.3), (4.8), and (4.16), the model described above can be used to determine the unknown geometric parameters of the structure. As a case study,

Table 4.1 Geometry parameters for the mechanism

Parameter (Symbol)	Value
Length of flexible links 2 and 4 (l_2, l_4)	36 mm
Width of flexible links 2 and 4 (c_2, c_4)	10 mm
Thickness of flexible links 2 and 4 (h_2, h_4)	0.5 mm
Length of flexible link 3 (l_3)	30 mm
Width of flexible link 3 (c_3)	10 mm
Thickness of flexible link 3 (h_3)	3 mm
Length of flexible link b_3 (l_{b3})	36 mm
Width of flexible link b_3 (c_{b3})	10 mm
Thickness of flexible link b_3 (h_{b3})	6 mm
Initial angle of links 2 and 4	75°

the chosen geometry parameters which satisfy the design criteria are summarized in Table 4.1. The flexure parallel linear guiding stage is designed with the main parameters: $l_2 = l_4 = 38$ mm, $b_3 = 36$ mm, $h = 0.5$ mm, and $c = 10$ mm.

4.4.1 Static Performance Analysis with FEA

The quantitative models are developed under the ideal situations. In order to test the performance of the designed stage and to verify the accuracy of the derived models, FEA simulation is conducted by using ANSYS software package. The material of the stage is selected as Al-7075 alloy, which is easy to machine, and its main characteristic parameters are: Young's modulus = 71.7 GPa, yield strength = 503 MPa, Poisson's ratio = 0.33, and density = 2.81×10^3 kg/m^3.

The quantitative models as developed in previous section predict that, the work range of the XY stage in each direction is ±7 mm. The deformation result is shown in Fig. 4.8b. In addition, Eq. (4.9) evaluates that the maximum force required to drive the stage is $F_{\max} = 180.5$ N. However, the FEA simulation result shows that the required actuation force is 160 N. Taking the FEA result as the benchmark, it is seen that the established quantitative models predict the stage performance with a deviation of 12.8 % with respect to FEA result.

Moreover, when one input displacement arrives at 7 mm, the deformation result of FEA static simulation is illustrated in Fig. 4.8a. It shows the compliance of the Y-axis direction when . The yield strength of the Al-7075 is $[\sigma] = 503$ MPa. As can be seen from Fig. 4.9, the induced maximum stress is 485.25 MPa $< [\sigma] =$ 503 MPa, which shows a safety factor of 1.27 of the material, meeting the strength condition.

In addition, Fig. 4.10 shows the compliance of the Y-axis direction when $\delta x = 0$. As compared with the FEA result, the result of the developed quantitative model is 5.6 % lower. The discrepancy mainly comes from the small deformation assumptions of the PRBM.

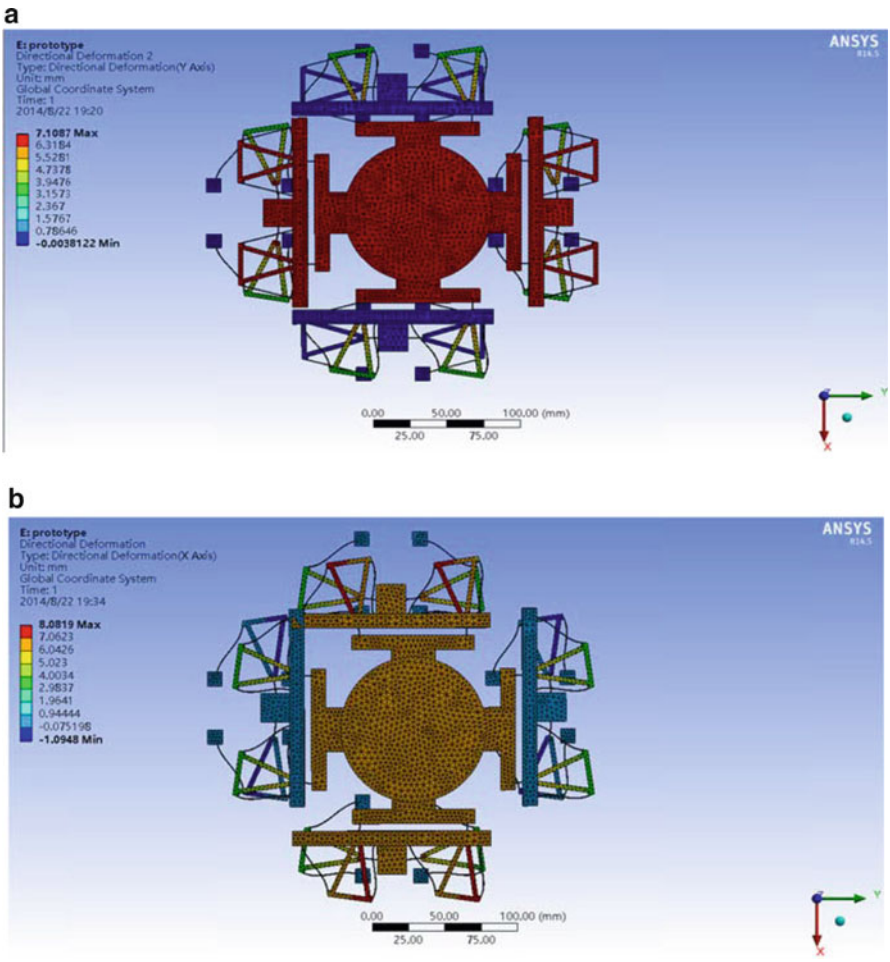


Fig. 4.8 Static FEA results of the XY stage under different conditions. (**a**) Deformation result when $\delta_y = 7$ mm and $\delta_x = 0$ mm; (**b**) deformation result when $\delta_y = 7$ mm and $\delta_x = 7$ mm

On the other hand, the FEA results as illustrated in Fig. 4.11 show the parasitic motion in Y-axis, which is induced by the actuation in X-axis. Figure 4.11a exhibits the FEA results of the symmetrical and asymmetric structure designs. It is obvious to observe that as compared with the asymmetric design, the symmetrical structure design can significantly reduce the parasitic motion in the non-working direction by over 60 times.

The new Roberts mechanisms can produce a straight-line motion under the working situation. As shown in Fig. 4.11 and Table 4.2, when the stage delivers the full work range in the working direction (i.e., 7 mm in X-axis), the central point's parasitic motion in the non-working direction of the output stage (i.e., Y-axis) is 0.64 μm. It means that the maximum cross-axis displacement caused by the work direction is only 0.009 % in theory.

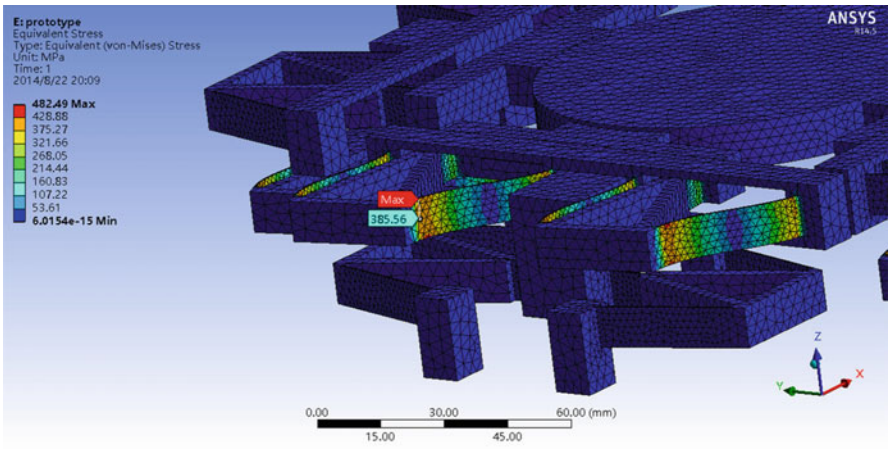


Fig. 4.9 FEA results of stress distribution

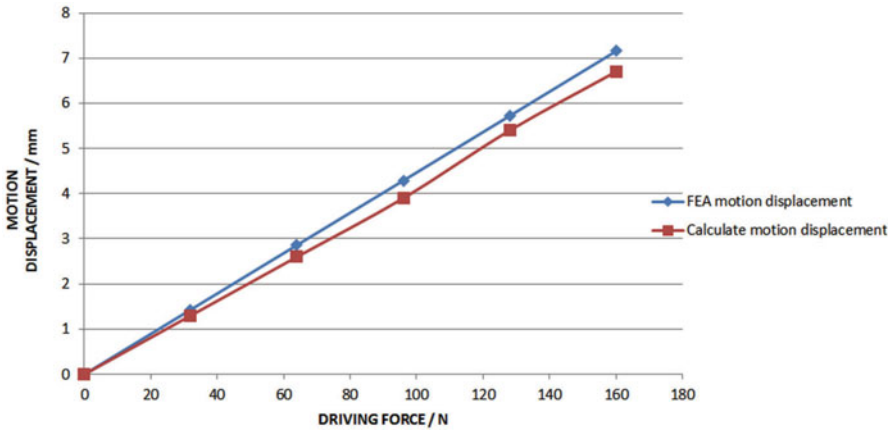


Fig. 4.10 The motion displacement comparison between FEA and calculated results

4.4.2 Dynamic Performance Analysis with FEA

The FEA results of dynamic performance analysis are shown in Fig. 4.12. It is observed that the first-three natural frequencies are 32.3 Hz, 32.5 Hz, and 58.7 Hz, respectively. As described in the previous section, the designed natural frequency is 40 Hz. This reveals that the developed quantitative dynamic model predicts the stage's natural frequency with a deviation of 18.75 % with respect to FEA result.

Besides, the modal shapes of the first two modes are shown in Fig. 4.12a, b, which are the translations along the two working axes. The third one as shown in Fig. 4.12c is the rotation of the output platform in the working plane. The first six resonant frequencies are shown in Table 4.3. The similar values of the first

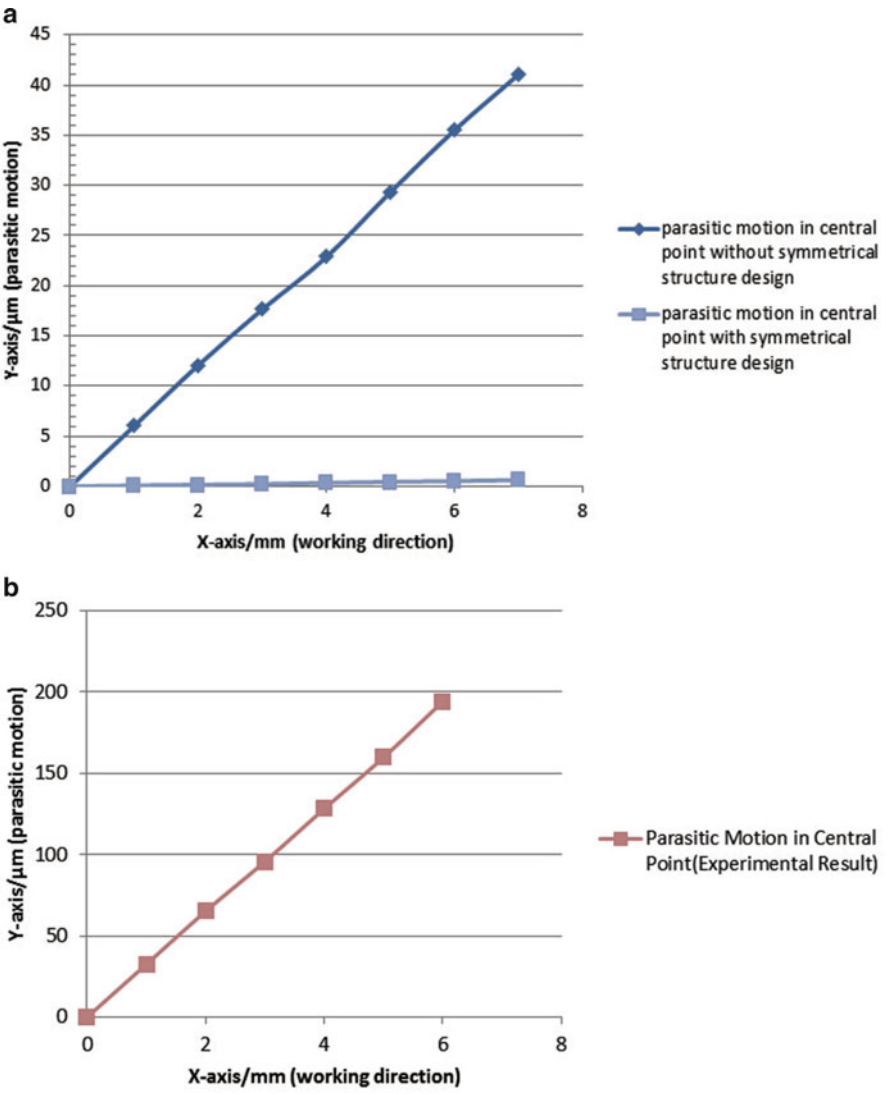


Fig. 4.11 Comparison of parasitic motions for the XY stage. (**a**) FEA simulation results of parasitic motions of the XY stage with and without symmetrical structure design; (**b**) experimental results of parasitic motions of the XY stage with symmetrical structure design

two natural frequencies indicate that the XY stage has almost the same dynamics performance in the two working directions, which benefits from the parallel-kinematic design of the stage. In addition, the frequency of the rotational mode is almost twice higher than the first-two modes, which confirms that the stage has two desired degrees of freedom.

Table 4.2 FEA simulation results of the parasitic motion with and without symmetrical structure design

Working displacement in X-axis (mm)	The parasitic motion in Y-axis without symmetrical structure design (μm)	The parasitic motion in Y-axis with symmetrical structure design (μm)
0	0	0
1	6.0361	0.0839
2	12.091	0.1551
3	17.656	0.2297
4	22.879	0.3552
5	29.362	0.4517
6	35.529	0.5429
7	41.076	0.6371

4.5 Prototype Fabrication and Experimental Study

4.5.1 Prototype Fabrication

The assembled CAD model of the designed XY stage is shown in Fig. 4.13a. The XY stage has been fabricated using Al-7075 alloy through the wire electrical discharge machining (WEDM) process. The prototype is shown in Fig. 4.13b, which is driven by two voice coil motors (VCMs). The output positions are measured by two Keyence laser displacement sensors. The controller is constructed by NI CRIO 9022 and 9018 FPGA module. Control algorithm is realized by NI LABVIEW software package.

Since the FEA result reveals that $F_{\max} = 160\,\text{N}$, two VCM motors (model: NCC05-18-060-2X, from H2W Techniques, Inc.) are selected to drive the stage. They can provide a sufficient large output force of 194.6 N with a stroke of 12.7 mm. The total dimension of the stage is $244 \times 244\,\text{mm}^2$. So, the area ratio of the stage is 0.32 %. It is seen that the proposed stage exhibits a relatively large area ratio, which means that the space is fully used by the two-layer design.

The two-layer design enables the generation of relatively compact physical size in comparison with a monolithic design. However, the shortcoming of two-layer design lies in that assembly is needed to develop the XY stage. Considering that the fixing holes can be manufactured with a high precision of $\pm 5\,\mu\text{m}$, a low level of assembly error is expected. The performance of the developed XY stage is tested through the subsequent experimental studies.

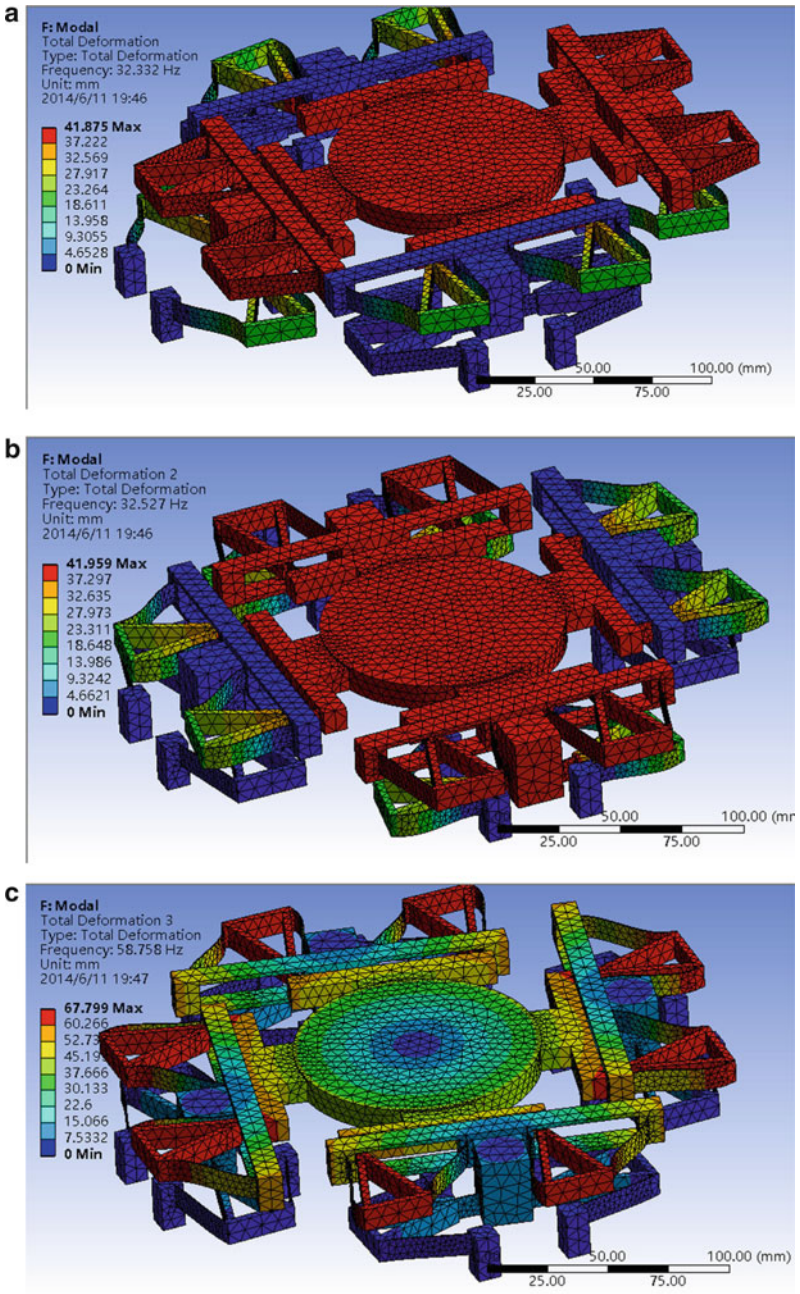


Fig. 4.12 FEA simulation results of modal analysis. (**a**) The first resonant mode; (**b**) the second resonant mode; (**c**) the third resonant mode

Table 4.3 The first six frequencies of resonant modes

Mode	Frequency (Hz)
1	32.33
2	32.53
3	58.76
4	212.74
5	324.33
6	326.42

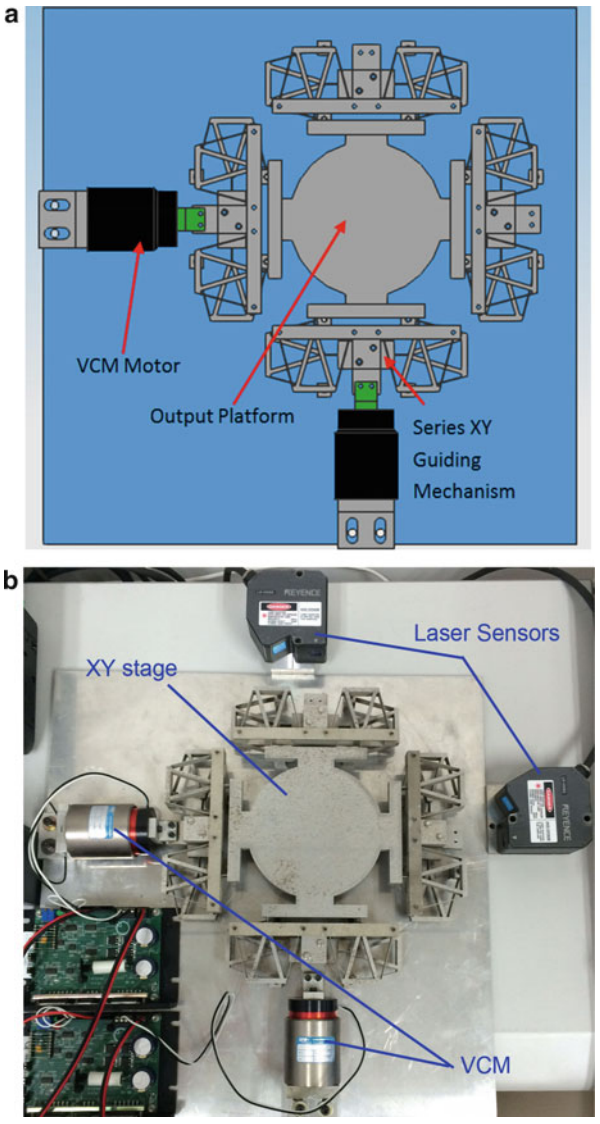


Fig. 4.13 (**a**) CAD model and (**b**) fabricated prototype of the XY micropositioning stage

4.5.2 Experimental Investigations

In this section, experimental studies are carried out to test the static and dynamic performance of the developed XY micropositioning stage.

First, the stroke of the XY stage is tested by applying a 0.1-Hz, 7-V sinusoidal voltage signal to drive the each VCM at a time. For instance, by applying the signal to the *X*-axis VCM as shown in Fig. 4.14a, the output displacement in *X* and *Y* axes are shown in Fig. 4.14b, c, respectively.

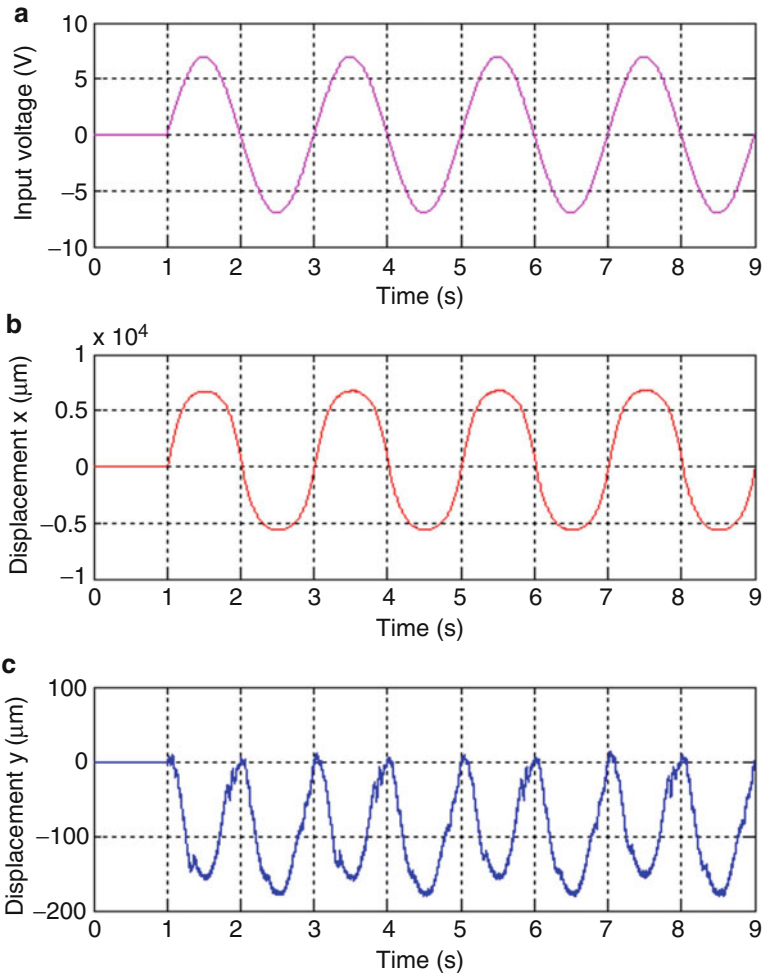


Fig. 4.14 Experimental results of *X*-axis motion range testing. (**a**) Voltage applied to VCM; (**b**) output displacement in *X*-axis; (**c**) output displacement in *Y*-axis

It is seen from Fig. 4.14 that the motion range in X-axis is 12.4 mm, and the caused coupling motion amounts to 195.0 μm. This indicates a crosstalk of $\frac{195\,\mu m}{12.4\,mm} \times 100\,\% = 1.57\,\%$ between the axes. Similar results are produced when the VCM in Y-axis is actuated. That is, the motion range in Y-axis is 12.2 mm, and the coupling motion in X-axis is 202.3 μm, which exhibits a crosstalk of $\frac{202.3\,\mu m}{12.2\,mm} \times 100\,\% = 1.66\,\%$ between the two working axes.

Due to the manufacturing error and assembly error of the XY stage, the experimental results (1.57 % and 1.66 %) of crosstalk is much bigger than the FEA simulation result (0.09 %). If the machining and assembly precision can be improved, the discrepancy between the simulation result and experiment result will be reduced. In practice, to reduce the crosstalk values and to achieve precise positioning, suitable feedback control algorithm can be implemented to cooperatively control the output displacements in the two working axes of the XY micropositioning stage.

Next, the dynamic performance of the XY stage is examined by applying a swept-sine voltage signal to drive each VCM alternately. By driving the X-axis VCM using a swept-sine voltage with the frequency ranging from 0.1 to 1000 Hz, the produced X-axis motion response is recorded by the laser displacement sensor. The frequency response is obtained by FFT algorithm and the result is shown in Fig. 4.15. It is found that the stage exhibits a natural frequency of 18.6 Hz in the X-axis. Moreover, the natural frequency in the Y-axis is obtained as 19.5 Hz, which is similar to the X-axis value.

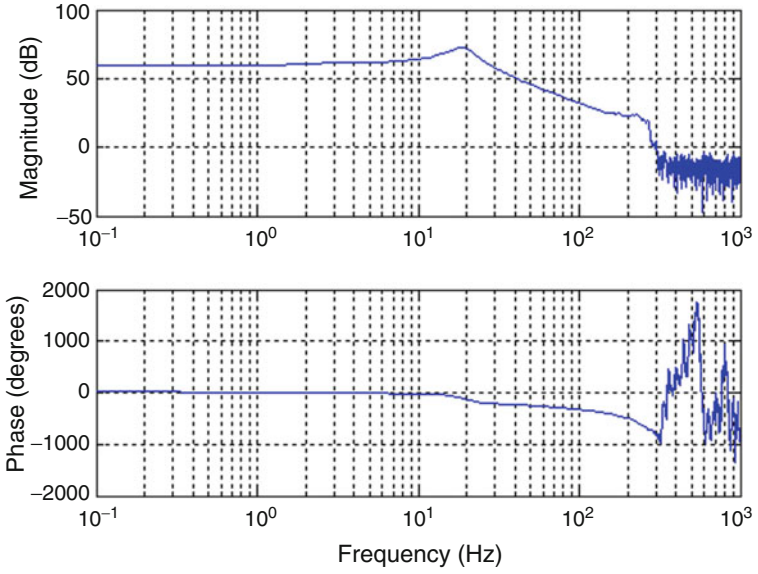


Fig. 4.15 Frequency response of the XY stage in X-axis direction

4.5.3 Experimental Study with PID Control

To demonstrate the positioning performance of the XY stage, closed-loop control is implemented using PID control. For the purpose of designing the PID controller of the stage, only the x-axis motion control is handled in this research, because the y-axis motion is similar to x-axis performance.

The input/output datasets are used to identify the plant transfer function by estimating the model from the frequency response data. The frequency responses obtained by experimental study are shown in Fig. 4.15. The first resonant mode occurs around 16.6 Hz. According to the Bode plot as shown in Fig. 4.15, a second-order plant model is identified as shown below.

$$G(s) = \frac{5s^2 - 1.8 \times 10^4 s + 2 \times 10^7}{1.2s^2 + 36s + 2 \times 10^4}. \tag{4.24}$$

The model matches the system dynamics well in the frequencies up to 100 Hz. In order to capture high frequency dynamics accurately, a much higher order model is required to be identified.

4.5.3.1 Simulation Study

According to the transfer function (4.24), a PID control is realized in MATLAB and Simulink, as shown in Fig. 4.16. The PID control parameters are adjusted according to the PID Tuner in the MATLAB/Simulink. In order to get a combination property of the PID control, according to automatic adjustment of the control parameters, the manual tuning is also applied. Then, the PID control parameters K_p, K_i, K_d, and filter coefficient (N) are 0.006, 37.282, 0.006, and 730337.058, respectively.

The purpose of the control simulation is to test the feasibility of the PID controller. Figures 4.17 and 4.18 show the system responses to a step input and a sinusoidal input, respectively, where the dashed lines represent the reference inputs and the solid lines represent the output responses. It can be seen that the system

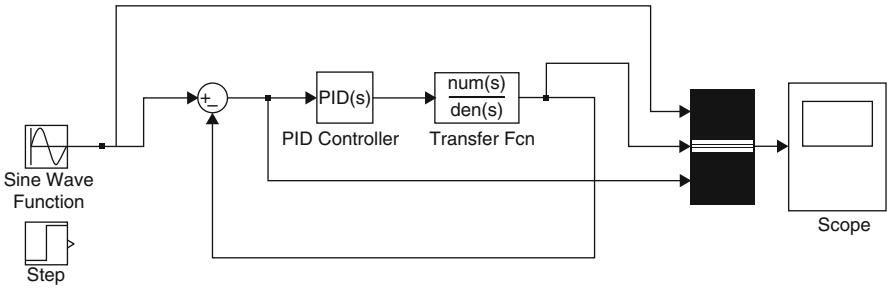


Fig. 4.16 MATLAB simulation model with PID control

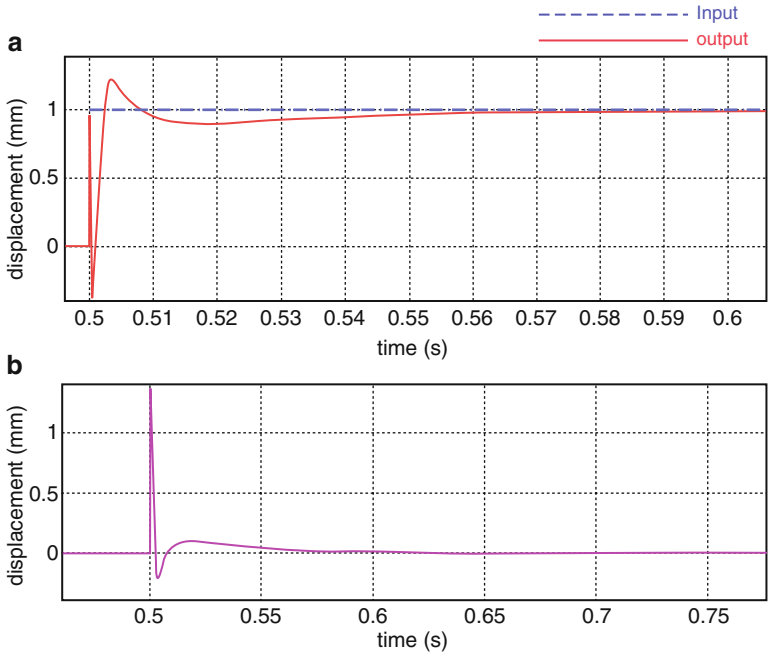


Fig. 4.17 Simulation results of step response with PID control. (**a**) Step input signal response; (**b**) step response error

performance is not very good, but the steady state error is small as shown in both cases. The simulation results are shown in Table 4.4.

4.5.3.2 Experimental Study

After the simulation study with the PID control, experimental study is carried out on the prototype XY stage. The system is driven by two VCMs. The output positions are measured by two Keyence laser displacement sensors. Control algorithm is realized by NI LABVIEW software package. The diagrams of the control algorithm program are shown in Fig. 4.19.

In order to test the performance of the system, different types of signal are chosen as the input signal of the PID control system. The sampling frequency of the laser sensor is set as 1000 Hz. To remove the noises of sensor output signal, a second-order low-pass filter with a cut-off frequency of 5 Hz is adopted in data acquisition process.

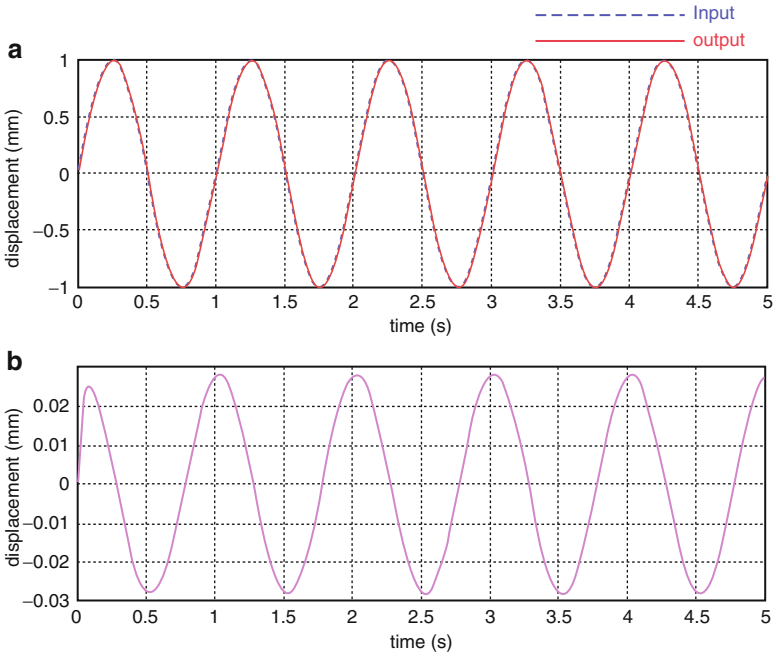


Fig. 4.18 Simulation results of sinusoidal tracking with PID control. (**a**) Sinusoidal tracking result; (**b**) sinusoidal tracking error

Table 4.4 Simulation results of the PID controller

	Step response		Sinusoidal response
Maximum overshoot (%)	2 % settling time (s)	Steady-state error (mm)	Steady-state error bound (mm)
16	0.4	0.002	±0.028

The experimental result of step response is shown in Fig. 4.20. In addition, Fig. 4.21a–c show the system responses to 0.1-Hz, 0.5-Hz, and 1-Hz sinusoidal inputs, respectively, where the blue dashed lines represent the reference inputs and the red solid lines represent the output responses. A summary of experiment results with different inputs are shown in Table 4.5.

In view of Tables 4.4 and 4.5, it is observed that as compared with the simulation result, the experimental study produce similar trend of results. However, the experimental results exhibit a time delay of about 50 ms. With sinusoidal reference input, the delay time is proportional to the frequency of the input signal. This time delay is mainly induced by the bandwidth of the control system. In addition, the delay may also arise from the low-pass filter (5-Hz cut-off frequency) which is used in the data acquisition process.

In order to identify the positioning accuracy, one part of the stable error signal has been extracted. This part of signal has 1000 sample points. Then, it is used to

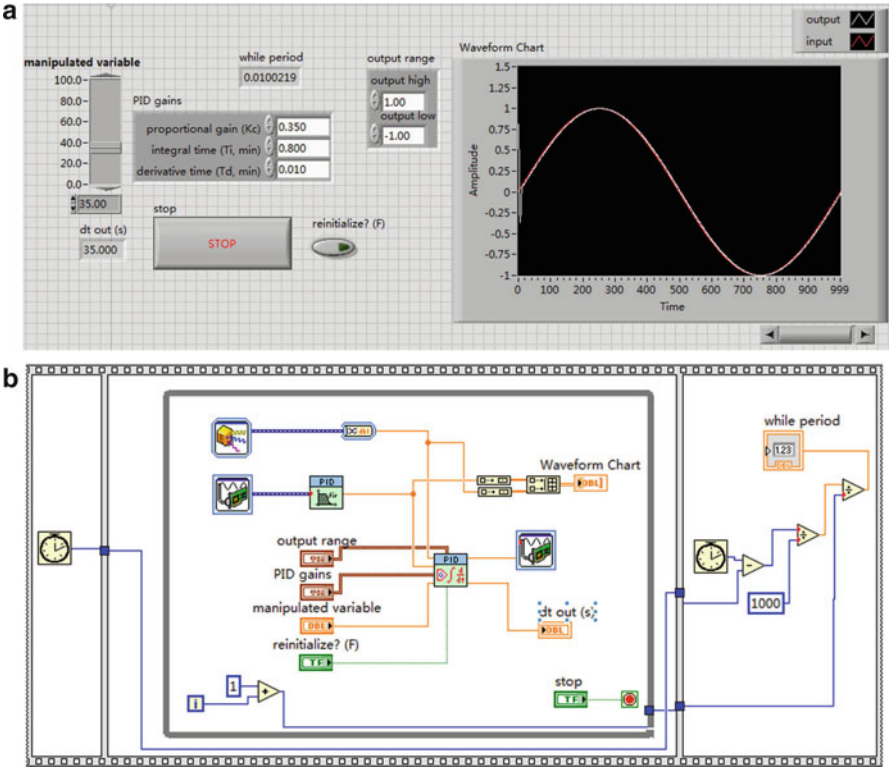


Fig. 4.19 NI LABVIEW control algorithm program. (**a**) Front panel; (**b**) program diagram

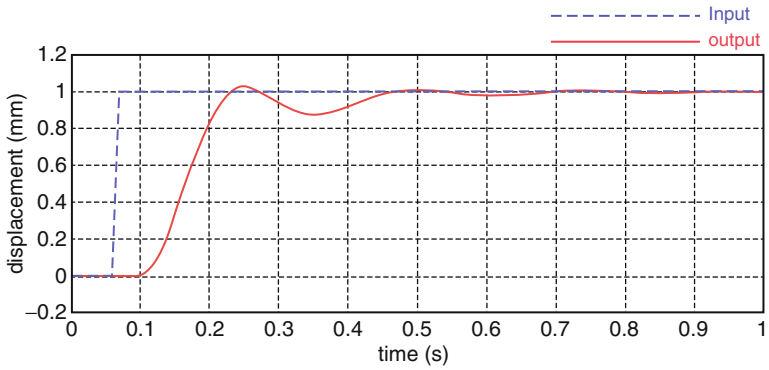


Fig. 4.20 Experimental results of step response with PID control

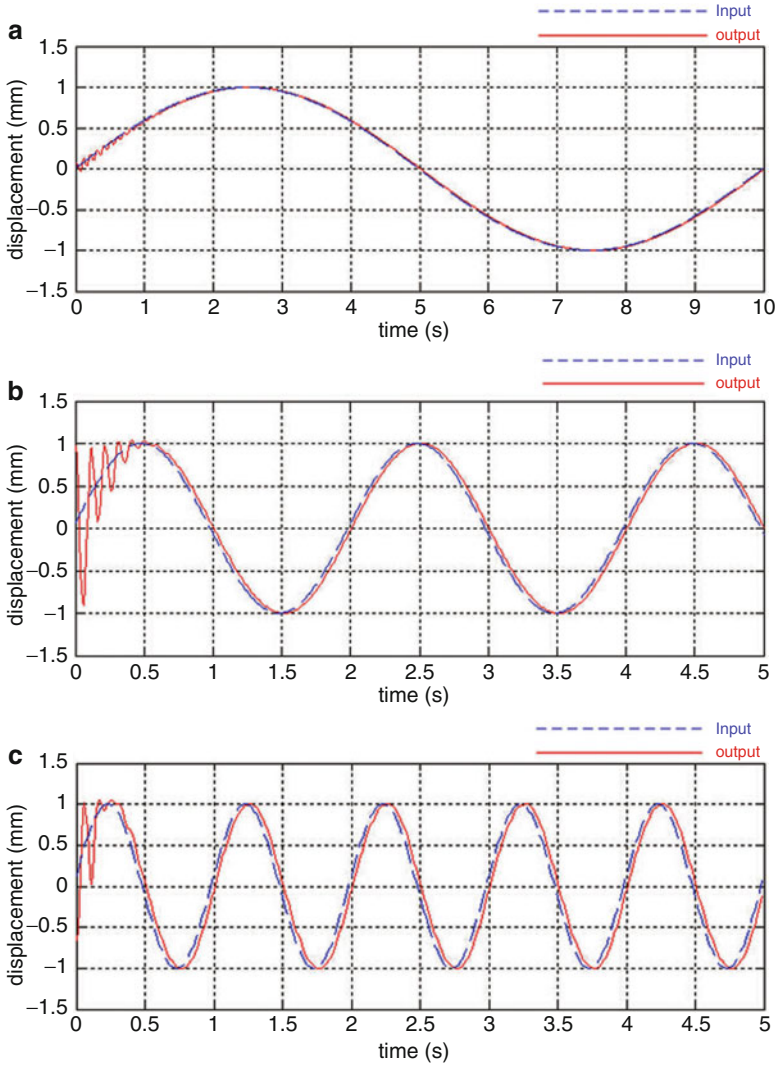


Fig. 4.21 Experimental results of sinusoidal tracking with PID control. (**a**) 0.1-Hz sinusoidal input; (**b**) 0.5-Hz sinusoidal input; (**c**) 1-Hz sinusoidal input

fit a normal distribution curve as shown in Fig. 4.22. According to the ISO standard as shown in (4.25), the positioning accuracy should be $\pm 3\sigma$. So, the positioning accuracy of the stage is $2\times 3\sigma = 0.96\,\mu\text{m} \approx 1\,\mu\text{m}$.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \text{ where : } n = 1000. \tag{4.25}$$

Table 4.5 Experiment results of the PID control system

	Step response		
Type of signal	Maximum overshoot (%)	2 % settling time (s)	Steady-state error (mm)
Step input	2.5	0.46	0.003
	Sinusoidal response		
	Steady-state error bound (mm)	Mean absolute error (mm)	Root mean square error (mm)
0.1-Hz sine wave	± 0.02	0.012	0.015
0.5-Hz sine wave	± 0.1	0.067	0.107
1.0-Hz sine wave	± 0.2	0.130	0.156

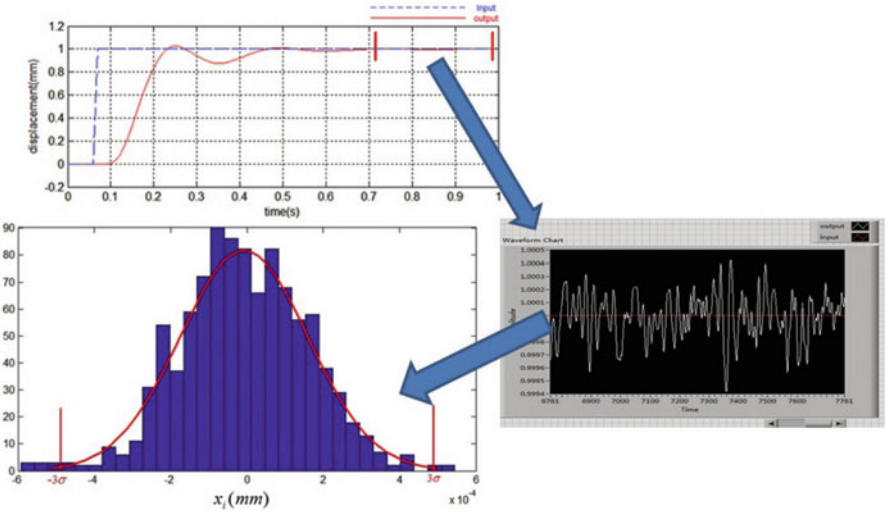


Fig. 4.22 Testing results of the stage positioning accuracy

4.5.4 Discussion on Stage Performance

The experimental results demonstrate a large stroke of the developed XY micro-positioning stage. The crosstalk between the two working axes may be induced by the fabrication errors of the stage parameters, and assembly errors of the stage and motors. In addition, due to the stroke limit of the adopted VCM motors, the experimental result of the XY stage stroke is slightly lower than the analytical result. The desired range of ±7 mm can be achieved by employing VCMs with larger stroke.

The experimental testing reveals that the XY stage exhibits the natural frequencies of 18.6 and 19.5 Hz in *X*- and *Y*-axes, respectively. Such values are smaller than the FEA results of 32.3 and 32.5 Hz in the two working axes, respectively. These discrepancies are mainly caused by the mass of the moving coil of VCM motor, which is not considered in FEA simulation. In the future, a higher natural frequency

will be achieved by implementing an optimal design of the XY stage parameters. In addition, control schemes will be implemented to achieve a precise positioning control of the XY stage.

4.6 Conclusion

A new type of parallel-kinematic XY micropositioning stage based on flexure Roberts mechanism has been designed, analyzed, fabricated, and tested in this paper. FEA simulation results show that this kind of compliant mechanism not only has good performance in reducing the parasitic motion in non-working direction, but also can produce a large workspace of $14 \times 14\,\text{mm}^2$. So, it meets the requirements of a large stroke, high precision linear guiding mechanism. The large stroke and small crosstalk have been verified by experimental studies. Experimental testing demonstrates a natural frequency around 20 Hz in both working axes. Moreover, it owns a relatively compact structure thanks to a two-layer mechanism design. It is found that it is feasible to devise a large-stroke micropositioning stage based on the Roberts mechanism. At last, a PID system was designed to control the X–Y motion of stage, and it get a high manipulate precision (the positioning accuracy is 1 μm). Future work involves architectural optimization and precision control of the XY stage for pertinent applications.

Acknowledgements The work was supported by the Macao Science and Technology Development Fund under Grant Nos.: 070/2012/A3 and 052/2014/A1, and the Research Committee of the University of Macau under Grant Nos.: MYRG083(L1-Y2)-FST12-XQS and MYRG078(Y1-L2)-FST13-XQS.

References

1. A.G. Ambekar, *Mechanism and Machine Theory* (Prentice-Hall of India, New Delhi, 2007)
2. S. Awtar, A.H. Slocum, Constraint-based design of parallel kinematic XY flexure mechanisms. J. Mech. Des. **129**(8), 816–830 (2007)
3. W. Chen, J. Jiang, W. Chen, J. Liu, A novel flexure-based uniaxial force sensor with large range and high resolution. Sci. China Technol. Sci. **58**(8), 1940–1948 (2013)
4. W. Chen, J. Jiang, J. Liu, S. Bai, W. Chen, A passive eddy current damper for vibration suppression of a force sensor. J. Phys. D Appl. Phys. **46**, 075001 (2013)
5. Y.J. Choi, S.V. Sreenivasan, B.J. Choi, Kinematic design of large displacement precision XY positioning stage by using cross strip flexure joints and over-constrained mechanism. Mech. Mach. Theory **43**(6), 724–737 (2008)
6. K.B. Choi, H.J. Lim, G.H. Kim, J.J. Lee, A flexure-based scanner for a fully bidirectional operation driven by a differential piezo force. Proc. IME C J. Mech. Eng. Sci. **228**(7), 3186–3199 (2014)
7. D. Devaprakasam, S.K. Biswasa, Design of a flexure for surface forces apparatus. Rev. Sci. Instrum. **74**(3), 1228–1235 (2003)
8. L.L. Howell, *Compliant Mechanisms* (Wiley, New York, 2001)

9. Z.B. Huang, W.J. Ge, L.E. Ma, Research status quo of flexible mechanism and its application prospect in the field of bionic. J. Mach. Des. Res. **20**(21), 279–252 (2004)
10. N.B. Hubbard, J.W. Wittwer, J.A. Kennedy, D.L. Wilcox, L.L. Howell, A novel fully compliant planar linear-motion mechanism, in *Proceedings of ASME 2004 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pp. 1–5, Salt Lake City (2004)
11. D. Kang, K. Kim, D. Kim, J. Shim, D.G. Gweon, J. Jeong, Optimal design of high precision XY-scanner with nanometer-level resolution and millimeter-level working range. Mechatronics **19**(4), 562–570 (2009)
12. Y. Lei, B.Y. Chen, Y.Z. Yang, Nanoscale research present situation and development trend of micro bench. J. Zhejiang Univ. Technol. **23**(1), 72–82 (2006)
13. Y.M. Li, Q.S. Xu, Modeling and performance evaluation of a flexure-based XY parallel micromanipulator. Mech. Mach. Theory **44**, 2127–2152 (2009)
14. Z. Li, J. Yu, A novel large-range XY compliant parallel micro-manipulator. J. Changzhou Univ. **25**(4), 14–19 (2013)
15. Y.T. Lin, J.J. Lee, Structural synthesis of compliant translational mechanisms, in *Proceedings of 12th IFToMM World Congress*, pp. 1–5 (2007)
16. S. Polit, J. Dong, Development of a high-bandwidth XY nanopositioning stage for high-rate micro-/nanomanufacturing. IEEE/ASME Trans. Mechatron. **16**(4), 724–733 (2011)
17. X.D. Shi, J.L. Li, L.W. Wang, Design of large displacement and high precision linear stage based on laminated flexure hinge and ultrasonic motor. Manuf. Technol. Mach. Tool **2**, 104–106 (2007)
18. K. Tan, G.H. Zong, S.S. Bi, Z. Yu, Multi-leaf configuration of expanded-motion-range flexure hinges. ProcDual Use Technologies & Products **3**, 38–39 (2007)
19. X. Tang, I.M. Chen, Q. Li, Design and nonlinear modeling of a large-displacement xyz flexure parallel mechanism with decoupled kinematic structure. Rev. Sci. Instrum. **77**(11), 115101 (2006)
20. B.P. Trease, Y.M. Moon, S. Kota, Design of large displacement compliant joints. J. Mech. Des. **127**(4), 788–798 (2004)
21. H. Wang, X.M. Zhang, Input coupling analysis and optimal design of a 3-DOF compliant micro-positioning stage. Mech. Mach. Theory **43**, 400–401 (2008)
22. X.K. Wang, D. Wu, C.Y. Cheng, J. Mech. Eng. **10**(5), 570–576 (1999) (Chinese)
23. Wikipedia (2007) http://commons.wikimedia.org/wiki/File:Roberts_linkage.gif
24. Q.S. Xu, New flexure parallel-kinematic micro-positioning system with large workspace. IEEE Trans. Robot. **28**(2), 478–491 (2012)
25. Q.S. Xu, Design, testing and precision control of a novel long-stroke flexure micropositioning system. Mech. Mach. Theory **70**, 209–224 (2013)
26. Q.S. Xu, Design and development of a compact flexure-based XY precision positioning system with centimeter range. IEEE Trans. Ind. Electron. **61**(2), 893–903 (2014)
27. Q.S. Xu, Y.M. Li, Analytical modeling, optimization and testing of a compound bridge-type compliant displacement amplifier. Mech. Mach. Theory **46**, 183–200 (2011)
28. S.L. Xu, D. Chen, Precision ultra-precision positioning technology and its application. J. Mech. Eng. **8**(4), 73–75 (1997) (Chinese)
29. Y.D. Yan, T. Sun, S. Dong, With new drive micro displacement worktable design and experimental research. J. Mech. Eng. **37**(5), 5 (2001)
30. Y.K. Yong, T.F. Lu, Relationship among input-force, payload, stiffness and displacement of a 3-DOF perpendicular parallel micro-manipulator. Mech. Mach. Theory **45**, 756–771 (2010)
31. Y.Q. Yu, L.L. Howell, C. Lusk, Y. Yue, M.G. He, Dynamic modeling of compliant mechanisms based on the pseudo-rigid-body model. J. Mech. Des. **127**, 760–765 (2005)
32. Z.J. Yuan, X.K. Wang, *Precision and Ultra-Precision Machining Technology*, pp. 1–3. (Machinery Industry Press, Beijing, China, 1999)
33. H. Zhao, Z.Q. Lv, The present situation and development trend of piezoelectric micro-positioning stage. J. Mod. Mach. **4**, 33–35, 52 (2001)
34. S.S. Zhao, S.S. Bi, G.H. Zong, J. Yu, New large-deflection flexure pivot based on curved flexure element. Chin. J. Mech. Eng. **45**(4), 8–12 (2009)

Chapter 5
Electro-Magnetic Nano-Positioning

Zhuo Zhang, Yueqing Yu, Xinyu Liu, and Xuping Zhang

Abstract Electro-magnetic actuation technology has the potential of providing millimeters of actuation range with nanometers of positioning resolution, and recently has received increasing attention in the field of nano-positioning. Compared with other types of nano-positioning systems that are driven by electrostatic, thermal, and piezoelectric actuators, the electro-magnetic nano-positioning systems are able to achieve large power density with low input voltages, are capable of operating in large distances with virtues of linear response, and are featured with multi-axis actuation capability and with high bandwidth. Therefore, a variety of electro-magnetic nano-positioning systems have been developed for various nano-scale applications. Although electro-magnetic nano-position systems are equivalent in positioning objects in nano-scale, they could be different in size, configuration, degrees of freedom, fabrication process, and control scheme for being suited to certain application requirements. However, electro-magnetic nano-positioning systems generally consist of three essential parts: an electro-magnetic actuator, a guiding mechanism, and a positioning stage. Instead of covering all the aspects on either the various types of systems or the applications, this chapter presents the common and general principles of electro-magnetic nano-positioning systems: actuation, guiding mechanisms, kinematics, dynamics, control, and fabrication. The typical systems and applications in the literature are employed as examples to illustrate the common principles. The aim of the chapter is to provide readers with insight and guidance on the development of electro-magnetic nano-positioners.

5.1 Overview

Compared with other actuation types available such as electrostatic, thermal, and piezoelectric, etc., electro-magnetic actuations require low operation voltages, moderate power density, large operating distances, linear response, multi-axis capability, and high bandwidth [1]. These advantages make electro-magnetic nano-

Z. Zhang • Y. Yu • X. Liu • X. Zhang (✉)
Department of Engineering and Aarhus School of Engineering, Aarhus University,
Inge Lehmanns Gade 10, 8000 Aarhus C, Denmark
e-mail: xuzh@ase.au.dk

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_5

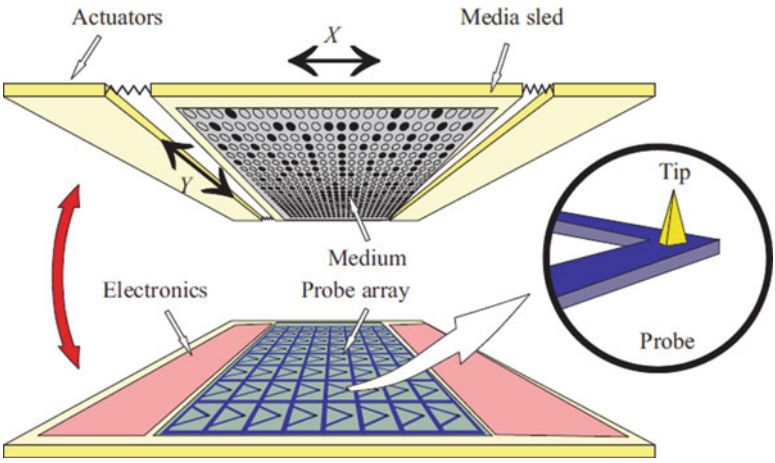


Fig. 5.1 Schematic diagram of the principle of probe-based data storage [6]

positioning systems increasingly attractive over the recent decades. There are emerging applications and techniques that could benefit from nano-positioning technology, such as the scanning probe microscopy (SPM) [2], probe-based data storage (PDS) [3, 4], and nano-scale electro-machining (nano-EM) [5], etc. This chapter is devoted to an overall introduction of the fundamental principles and guidance on the analysis, design, control, and fabrication of the electro-magnetic nano-positioning technology, i.e., nano-positioning apparatus driven by electro-magnetic actuators.

One application is the PDS that has been proposed in the late 1990s. This technology has emerged from SPM, where a surface is scanned and manipulated at the atomic level to form topographic images [6, 7]. This data storage system is MEMS-based, which is comprised mainly of two separate layers, as shown in Fig. 5.1 [6]. The upper layer depicted in Fig. 5.1 is the medium layer, on which data will be recorded or removed, whereas the other layer carries an array of probes (or heads, usually several thousands) that is used to read/write/erase data via scanning and contact with the medium. There are two main scanning configurations for PDS, i.e., (a) probe-scanning, where the probe moving relative to a fixed medium layer, and (b) medium-scanning, where the medium layer is movable while the probe layer remains stationary. Usually the tip size of the probes is atomic-scale, resulting in the data bits are approximately 15 nm in diameter and the tracking pitch of 30 nm. A pit that was written into the surface and the original surface represent logical “1” or “0,” respectively. In order to achieve data storage with high performance such as ultrahigh storage density, high speed, and suited to portable applications, the nano-positioning system must meet a stringent set of requirements. For example, to ensure extremely high data throughputs, a sufficient scanning range, high bandwidth, as well as reliable nano-scale positioning accuracy are needed. In addition, the portable storage applications also require lower energy consumption and vibration resistant capabilities.

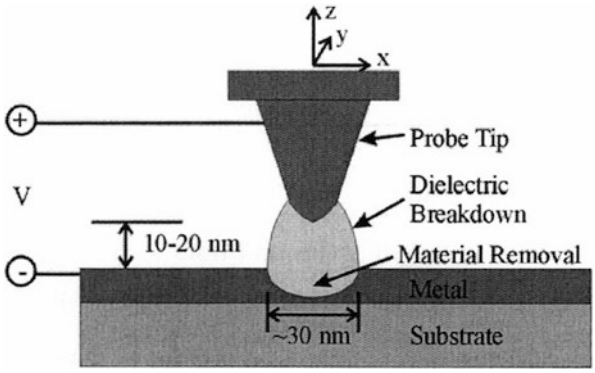


Fig. 5.2 Nano-scale electro-machining process [8]

Another important emerging application benefited from the nano-positioning technology is probe-based nano-scale electro-machining (nano-EM). This technique modifies a surface at nano-scale through removing conductive material from the work surface. The process is illustrated in Fig. 5.2 [8]. The gap between the probe and the work surface is filled with dielectric. When sufficient high electric potential is applied, the dielectric will breakdown resulting in removal of local material from the work surface. In nano-EM, the positioner must control the probe-tip location and orientation relative to the surface in all six degrees of freedom (DOFs), i.e., three translation DOFs: x, y, and z and three rotational DOFs: θ_x, θ_y, and θ_z. In addition to having the nanometer-level or better resolution, the nano-EM process also requires good repeatability and short execution time to make the technology practical and competitive.

Electro-magnetic nano-positioning technology enables us to move large or small objects in nanometer-level or better precision, however, the actual positioner is not necessarily be on nano-scale. The dimensions range from macro- [9] to meso- [8] and even micro-scales [10, 11] based on various application purposes. Macro-scale nano-positioners can easily operate over large distances (100s microns to centimeters) and exert large forces (N-kN) with nanometer precisions. One typical macro device is the magnetic levitation system [12–14], which achieve nanometer-level precision thanks to its non-contact nature. It eliminates mechanical friction, stiction, backlash, and hysteresis that are detrimental to precise control. In addition, the macro-sized system is well capable in integrating sensing and feedback systems in order to stabilize their motion and ensure nano-scale precision. However, the sensing and feedback control add complexity to the design and assembling of the system. The large mass leads to low natural frequencies and bandwidth. This limits applications where high throughput is needed. On the other hand, macro-scale nano-positioners often consume more energy, and exhibit more thermal growth that may cause thermal instability and errors. A macro six-axis nano-positioning device with magnetic levitation technology is shown in Fig. 5.3 [9].

Although macro-sized devices are capable of positioning objects in nano-scale, they are not amenable to massively parallel operation resulting in low throughput

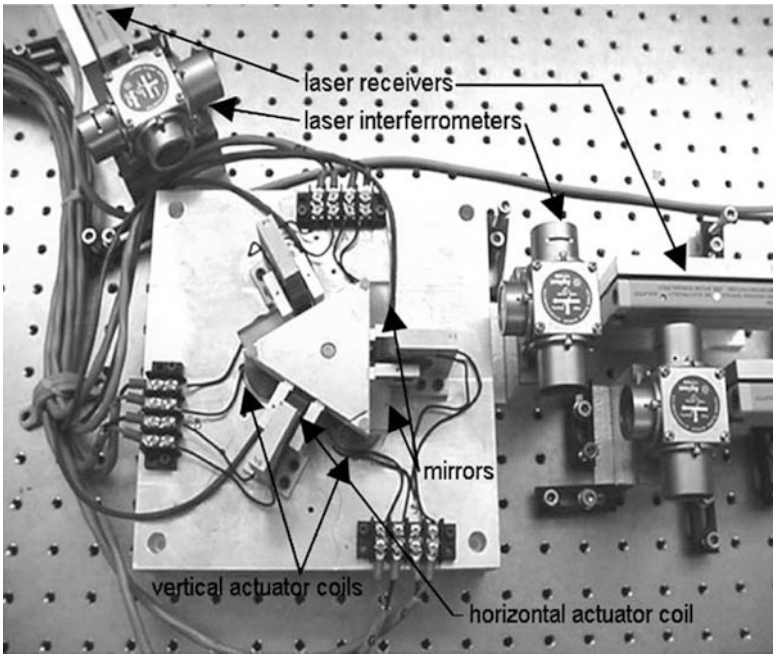


Fig. 5.3 Photograph of the six-axis maglev nano-positioning device [9]

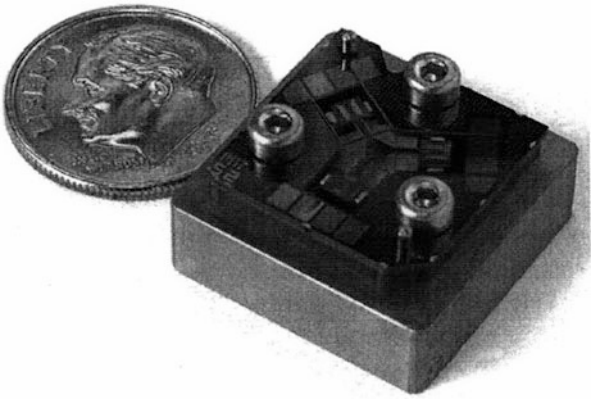


Fig. 5.4 Image of an *x*-axis meso-scale electro-magnetically driven nano-positioner developed by MIT. The dime is included for size reference [8]

rates. The great mass and volume limit the use for portable applications. Recent years, researchers from Massachusetts Institute of Technology (MIT) developed a six-axis meso-scale nano-positioner driven by electro-magnetic actuators, as shown in Fig. 5.4 [8]. This device is only centimeters in size with the central stage driven by three sets of integrated electro-magnetic actuators. Due to the reduced size, it is possible to run several nano-positioners in parallel to achieve high throughput rates. Moreover, the reduction in mass makes it much easier to attain higher mechanical

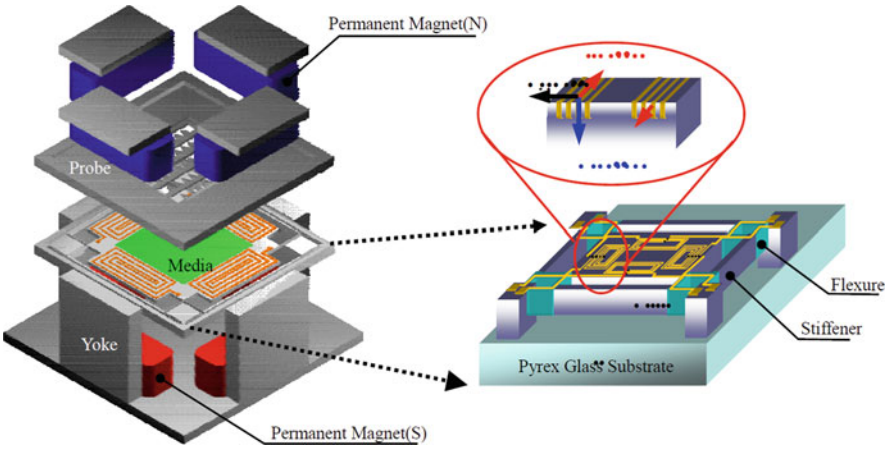


Fig. 5.5 Structure of probe-based data storage (PDS) with electro-magnetic micro x–y stage and its working principle [10]

bandwidth of the system. Also, meso-scale nano-positioners exhibit less thermal growth, cost less energy, and cause smaller thermally-induced errors accordingly than macro counterparts. However, several challenges must be overcome as the size is significantly scaled down. For example, design and integration of meso-scale multi-axis electro-magnetic actuators are totally different from macro devices. The assembling of position and orientation sensors is a big challenge.

With the maturation of MEMS (Micro-Electro-Mechanical System) technology, electro-magnetic nano-positioner can be scaled down to millimeter scale. Figure 5.5 shows a micro-scale nano-positioner for PDS application. The size of the media platform is 5×5 mm^2. The micro-scale positioner is compatible and portable. However, due to their small size it is difficult, if not impossible, to integrate sensors, and it is challenging to design a micro stage with up to six DOFs.

Despite their differences in size, the macro, meso, and micro electro-magnetic nano-positioner share a common operation principle and structure. The positioning system mainly consists of a stage, electro-magnetic actuation parts, and a guiding part that links the actuation parts to the stage. The purpose of this chapter is to introduce common principles on how electro-magnetic actuations could generate stage motions with nanometer precisions. Based on the principles and different applications, we will also point out key considerations on design, control, and fabrication. This chapter will not be confined to certain types of electro-magnetic nano-positioning systems, instead, it will provide potential readers with the basic principles, methodologies, and fundamental considerations on this topic.

5.2 Principle of Electro-Magnetic Actuation

5.2.1 Actuation Force Generation

In general, the electro-magnetic actuation parts consist mainly of one or several pairs of permanent magnets and coils. Upon the application of a current in the coil, the associated magnetic field of the coil is generated. In presence of the magnetic field generated by permanent magnets, the output actuation forces are generated thanks to the attraction and/or repel forces between coils and permanent magnets.

5.2.1.1 Solenoid

Solenoid is often used along with a permanent magnet to generate actuation forces based on magnetic attraction and/or repulsion effect. Solenoid behaves like a long bar magnet of which the north and south pole of the magnetic field is governed by right-hand rule, as shown in Fig. 5.6. There are a variety of applications benefited from solenoid such as electro-magnetic lifting machine, the electric bell, the electro-magnetic relay, and telephone earpiece, etc.

Based on Ampere's Circuital Law, the magnetic field of solenoid is expressed as

$$B = \mu_0 IN/L \tag{5.1}$$

where μ_0 is called the permeability of free space, I the input current, N the number of turns, i.e., number of loops the solenoid has, and L the length of solenoid. The strength of the magnetic field can be increased by the four approaches: (1) increasing the current; (2) increasing the number of turns; (3) placing a soft iron rod/core inside the solenoid; and (4) pushing the turns of the wire closer to make north and south poles closer. In practice, for positioning applications, the magnetic strength (actuation force) is governed by the control of input currents. In addition to the magnitude of the passing current, the direction of the current determines the direction of north and south poles and the direction of actuation accordingly.

The main shortcoming of the solenoid-based actuation is the difficulty to fabricate and assemble meso- and even micro-scale solenoids. They are used

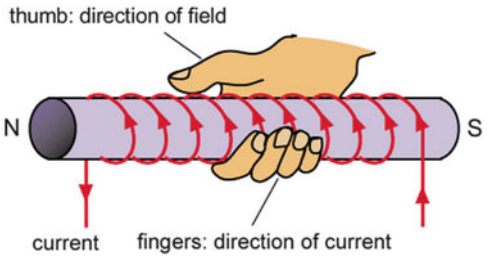


Fig. 5.6 Solenoid and right-hand rule

alone in macro-scale magnetic levitation systems for nano-positioning. Moreover, solenoid-based electro-magnetic actuators can only achieve one-axis motion, which limits their uses when multi-axis actuation is desired. These drawbacks are induced mainly by the spatial structure of the solenoid. To overcome this, planar-coil electro-magnetic actuation schemes have been employed in nano-positioning systems. To help reader better understand the actuation principles, the physical principle of Lorentz force and Ampere's force is briefly introduced as below.

5.2.1.2 Lorentz Force and Ampere's Force

In presence of only a magnetic flux $\mathbf{B}$, and electric charge q moving with a velocity $\mathbf{v}$ experiences a magnetic force $\mathbf{F_m}$. The magnetic force is called Lorentz force and is expressed as

$$\mathbf{F_m} = q\mathbf{v} \times \mathbf{B} \tag{5.2}$$

The magnetic force $\mathbf{F_m}$ affects only the direction of motion of q, not the kinetic energy. Because the direction of magnetic force is perpendicular to the velocity of q and does no work. Electric current is motion of charge. To make Lorentz force practical, Ampere's force has been employed which refers to the force exerts on the current-carrying wire in a magnetic field. The Ampere's force, total force acting on the wire, is

$$\mathbf{F} = \int_{\text{wire}} Id\mathbf{L} \times \mathbf{B} \tag{5.3}$$

where I represents electric current and $d\mathbf{L}$ refers to differential elements of length of the wire. Figure 5.5 shows a planar-coil-based electro-magnetic micro x–y stage, in which the Lorentz force/Ampere's forces used to actuate the stage. Usually, the magnetic field is created by permanent magnets, and thus the magnitude and direction of the magnetic force are determined by electric current.

5.2.2 Configurations

5.2.2.1 Magnet-Moving Versus Coil-Moving

In generating actuating force, two configurations of electro-magnetic actuator have been adopted: magnet-moving and coil-moving. Moving magnet actuations operates with the coil fixed and the magnets moves for actuation, whereas the moving coil systems works by fixing the magnets to the base and the coils are moved for actuation. A typical moving magnet system is shown in Fig. 5.7 [15], where the coils are fixed and the magnets are mounted on a flexure guiding mechanism. Fixing

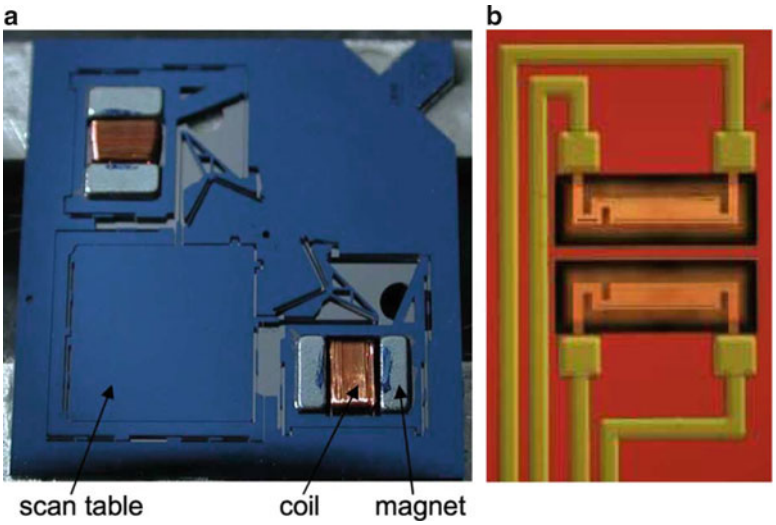


Fig. 5.7 (**a**) Photograph of a prototype nano-positioner for probe-based data storage and (**b**) the thermal position sensors [15]

the coils is ideal for thermal dissipation and therefore higher operating currents are allowed and larger actuation forces are achieved. However, the magnets must be bonded onto the flexure mechanisms, which is a challenging work especially for meso- and micro-scale devices. Another drawback inherent in magnet-moving design is that the magnets are in general a big mass which results in lower bandwidth of the actuators.

In contrast, for coil-moving configurations, as shown in Fig. 5.5, it is easier to achieve high bandwidth since the mass of coils can be decreased by virtue of micro-fabrications. However, for applications that require sufficient precision, the heat dissipation from the moving coil through the flexure mechanism will cause thermal-induced errors. To reduce the thermal-based errors, the operation current must be kept in lower magnitude for coil-moving configurations. As a result, such configurations can only provide less actuation force compared with magnet-moving configurations.

5.2.2.2 Single-Axis and Multi-Axis

Most of applications for nano-positioning require at least two degree-of-freedom, such as planar scanning. Some applications require up to six degree-of-freedom (DOF), such as nano-EM. For this purpose, electro-magnetic actuators must have one-axis or multi-axis capabilities, and need to be appropriately arranged. In general, for the purpose of achieving the positioning stage with multiple DOFs, either the choices of more actuators with less axes or less actuators with more

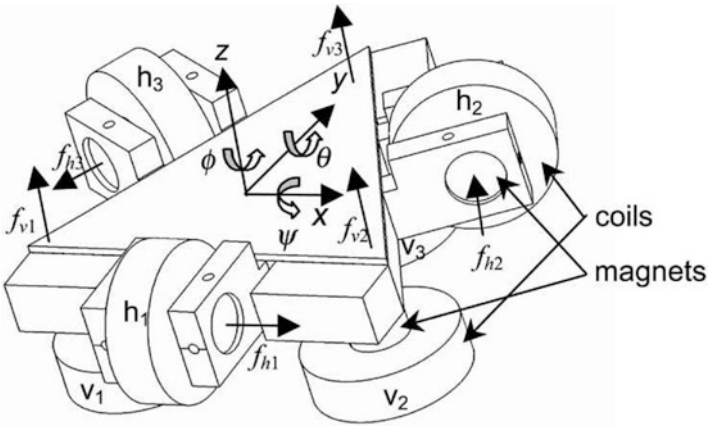


Fig. 5.8 A six DOFs nano-positioning device driven by six single-axis actuators [9]

axes is feasible. Fewer axes for each actuators will facilitate design of actuators with a compromise that the design of guiding mechanism and actuator arrangement become complex. On the contrary, more axes for each actuators may add complexity of actuator design but make design of guiding mechanism and actuator arrangement with ease. However, to the best knowledge of the authors, no efforts have been devoted to the development of electro-magnetic actuators having more than two-axis.

The six-axis magnetic levitation nano-positioning device shown in Fig. 5.3 has six independent single-axis electro-magnetic actuators, of which the detailed arrangement is seen in Fig. 5.8 [9]. In order to achieve six-DOF positioning of the stage, three actuators are combined to generate in-plane motion, and the rest of the three actuators are responsible in producing out-of-plane motions. Another six-DOF stage developed by MIT is shown in Fig. 5.4, where only three actuators with double-axis are incorporated. Each actuator provides two mutually perpendicular axial motions, i.e., in-plane and out-of-plane, as shown in Fig. 5.9. The three ways actuations simplified the guiding mechanisms that connect the actuators and the center stage. As a tradeoff, it adds complexity in design, fabrication, and assembling of each actuators.

As illustrated in Fig. 5.10 [16], in achieving both in-plane and out-of-plane motions, two separated coil layers are stacked. The stacked configuration is benefited from the planar coil-moving scheme. The actuation force is generated thanks to Lorentz force/Ampere's force via interaction of the coil current with the permanent magnet field.

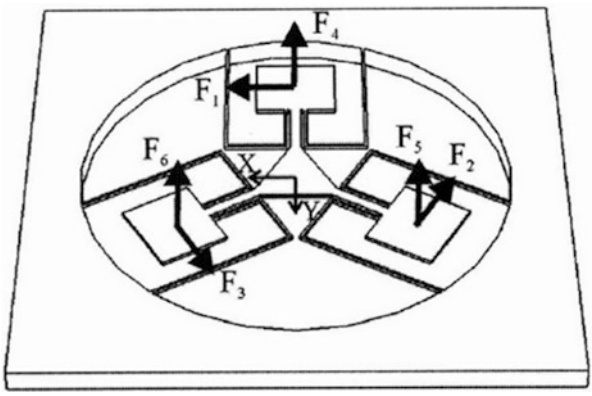


Fig. 5.9 A six DOFs nano-positioning device driven by three double-axis actuators [8]

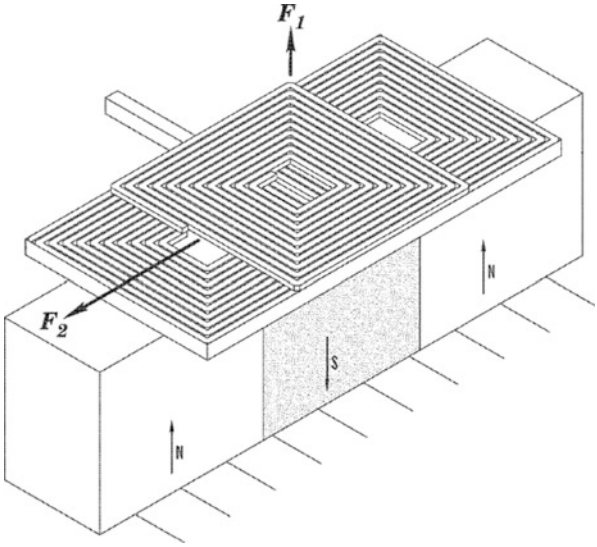


Fig. 5.10 Schematic diagram of a double-axis electro-magnetic actuator [16]

5.3 Compliant Guiding Mechanisms

Guiding mechanisms enable motion guidance of individual actuators to form the stage motion. There are two main categories of the guiding mechanisms: (1) the rigid mechanisms which use prismatic and rotational kinematic joints for bearing and guidance, as shown in Fig. 5.8; (2) the compliant mechanisms (i.e., flexure mechanisms) which transfer or transform motion, force, or energy via elastic deformations, as shown in Figs. 5.5 and 5.9. Compared with rigid guiding mechanisms, the compliant counterparts are more favorable particular in meso- and micro-scale applications. Since fabrication and assembling of rigid mechanisms are challenging and rigid mechanisms will incur backlash, inaccuracy in motion, frictions, and wear because of the rigid prismatic and rotational joints. In contrast,

the compliant mechanisms have advantages in two categories: cost reduction (part-count reduction, reduced assembly time, and simplified manufacturing processes) and increased performance (increased precision, increased reliability, reduced wear, reduced weight, and reduced maintenance) [17]. These virtues ideally fulfill the requirements as the size of electro-magnetic nano-positioner decreases. Moreover, this compliant structure enables stabilization without the feedback aids. There are also some challenges associated with compliant guiding mechanisms such as hysteresis, sensitivity to overloading, thermal sensitivity, and low range of motion, etc. In this section, we will provide readers with key methodologies and considerations in designing the compliant guiding mechanisms.

5.3.1 Design Considerations

Recent efforts in compliant guiding mechanisms are based on intuitions, for which the beams with constant sectional area are usually used. Upon operation and analysis, small and linear deformations are assumed. In designing the compliant guiding mechanisms for electro-magnetic nano-positioning systems, the finite-element analysis is used to obtain the force-deformation relationship, transmission ratio, thermal response, and stress, etc. In what follows, we list some key design considerations that can be used as a reference in achieving high nano-positioning performances for a variety of applications.

5.3.1.1 Distributed and Lumped Compliant Mechanisms

Distributed compliant mechanisms are those structures that they are flexible at every position, whereas the lumped compliant mechanisms are those consists of both rigid and flexible parts. For distributed compliant mechanisms the final position is determined by accumulation of deformations at every position. One typical geometry of distributed compliant mechanisms is the beam with uniform cross-sectional areas. As shown in Fig. 5.9, three sets of double-bent beams are used to support the stage and actuators and guide motions. In contrast, the final position of a lumped compliant mechanism is guided thanks to deformation of flexible hinges [18]. Different types of hinges provide unique properties to suit different requirements. Here we present three commonly used flexible hinge forms, i.e., the circular, the corner-filleted, and the elliptical hinges, as shown in Fig. 5.11 [19].

The circular hinge is superior to the others in achieving accurate rotational motions with high out-of-plane deflections. The center of deflection can be estimated to be at the center of the hinge. However, the high precision is only limited to a relatively small range. The corner-filleted hinge has relatively low bending stiffness. The performances of elliptical hinges are in between the circular and corner-filleted ones. For the same deflection, elliptical have lower maximum stress than corner-filleted ones resulting in a longer fatigue life. Besides the three

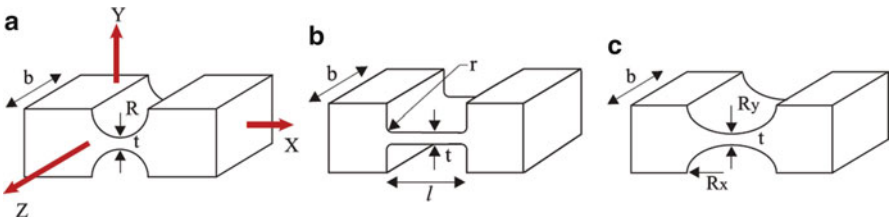


Fig. 5.11 Different types of hinges. (**a**) Circular, (**b**) corner-filleted, and (**c**) elliptical [19]

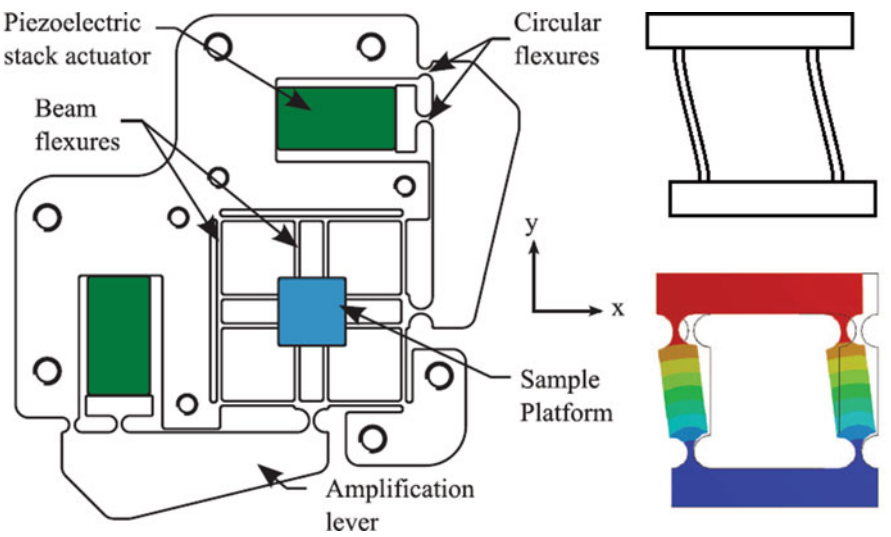


Fig. 5.12 Compliant parallel-guiding mechanisms in positioning application [19]

Table 5.1 A comparison of different flexure hinges [19]

Flexures	Stiffness		Motion accuracy	Max. stress
	In-plane	Out-of-plane		
Circular	High	High	High	High
Corner-filleted	Low	Low	Low	Low–med
Elliptical	Low–med	Low–med	Med–high	Low
Parabolic	Med–high	Med–high	Med–high	Med
Hyperbolic	High	High	High	Med–high

commonly used hinges, there are other flexure hinges available such as parabolic and hyperbolic. Details of the performances of those hinges are listed in Table 5.1 [19].

One typical compliant mechanism called compliant parallel-guiding mechanisms has been used in positioning applications, as shown in Fig. 5.12 [19], in which two opposing links remain parallel throughout the mechanisms motion. In this regard,

the parallel-guiding mechanisms produce alone the translation motion without a rotation motion exerting on the stage. To this end, the opposing links must be fabricated with the same length and form a parallelogram. This compliant four-bar mechanism can be either fully- or partially-compliant designed.

5.3.1.2 Stiffness and Resonant Frequency

The stiffness of the compliant guiding mechanisms determines the range of motion and resonant frequencies of the nano-positioner. For general cases, the compliant guiding mechanisms are used only confined to the linear elastic assumptions, where the positioning system can be treated as a linear mass-spring model. With this, the static displacement and resonant frequency of the system are given by [8]

$$F = k\delta \tag{5.4}$$

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \tag{5.5}$$

The static displacement and resonant frequency can be obtained for a value of stiffness. The stiffness of simple structures, such as uniform cross-sectional beams, can be derived using either Euler–Bernoulli theory or Castigliano's second theorem. For complex structures, finite-element methods are often used to obtain the stiffness. As can be seen in Eq. (5.4), for a force-limited actuation, a stiffer flexure leads to small range of motion. Equation (5.5) shows that for a given moving mass, a stiffer structure increases resonance frequency. Because the flexure structure deforms and stores potential energy, the resonance frequency will limit the bandwidth. As a result, the compliant guiding mechanisms should be designed neither too flexible nor too stiff. Another important concern is that the stiffness is sensitive to the fabrication errors, which may cause errors in the motion range and bandwidth.

5.3.1.3 Transmission Ratio and Coupling

The transmission ratio is defined as the ratio of the stage to actuation displacement in a given direction [8]. This concept is important in nano-positioning systems since structure of large transmission ratio will amplify the errors of the actuation causing large positioning errors in stage. The compliant guiding mechanisms may be designed to enable motion/error de-amplification for nano and even sub-nano-positioning applications. However, this de-amplification treatment will result in a limited range of motion.

For most cases, flexure-based systems will give rise to coupling problems, which is caused by non-zero stiffness of the flexures perpendicular to the direction of actuation. This will definitely change the transmission ratios and lead to additional parasitic motions.

5.3.1.4 Thermal Issues

For nano-motions, the thermal effects cannot be ignored [19]. As mentioned previously, for coil-moving configurations, the heat generates via Joule heating in the coils and dissipates through radiation, convection, and conduction. Among these, heat conduction to compliant guiding mechanisms predominates. The increased temperature will lead to thermal expansion effects, resulting in significant drift and positioning error. Careful material selection and mechanical arrangement are effective ways to minimize thermal expansion effects. Using the materials with small coefficient of thermal expansion (CTE) ensures less expansion with increased temperatures. Moreover, materials of high thermal conductivity should be selected in case that quick thermal equilibrium will be achieved. Proper arrangement of the compliant guiding mechanisms is essential in eliminating the thermal effects, which is presented in the following section.

5.3.1.5 Configurations

The configurations of compliant guiding mechanisms are considered at two aspects: function and performance. First, the configurations must fulfill the requirement that several actuations produce stage motions with certain number of DOFs. This is achieved by determining how many actuators with certain number of axes are used and how they should be allocated. Second, the arrangement of the compliant guiding mechanisms must ensure the positioning performance.

As mentioned previously, the thermal effects of the coil-moving actuator result in thermal drift and significant errors. In addition to a proper selection of materials, a symmetry structure is often employed for passive thermal compensation purpose, as shown in Fig. 5.13. In this scheme, the thermal drift and positioning errors are minimized because of the symmetric configuration. As a result, greater current can be applied on the coils and therefore the actuators can produce larger force without a loss of positioning error of the stage.

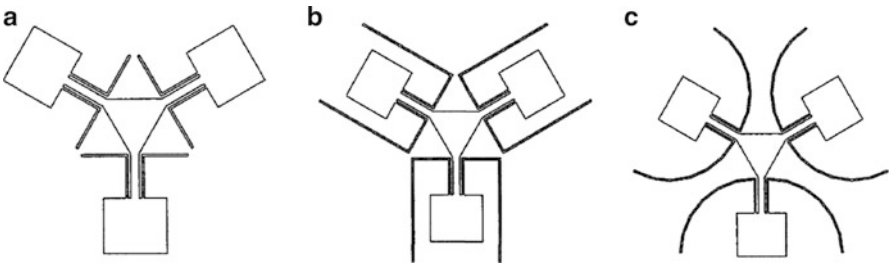


Fig. 5.13 Nano-positioner flexure concepts. (**a**) Single-bent beam, (**b**) double-bent beam, and (**c**) bow-flexure [8]

Figure 5.13 also illustrates three possible symmetric configurations, i.e., single-bent beam, double-bent beam, and bow-flexure [20, 21]. The geometric constraints limit the total length of the beams of the single-bent beam concept, which results in higher stiffness, higher resonant frequency, and lower transmission ratio over the other two concepts. The double-bent beam and bow-flexure share similar performances.

Compliant guiding mechanisms have the virtue of shock and vibration resistance because of their compliance. On the other hand, however, the compliant mechanisms must be designed stiffer to be robust against position errors caused by vibration and shock. To eliminate the shock and vibration effects on position errors and allow sufficient compliance for operations simultaneously, a mass balance concept is preferred. The nano-positioner in Fig. 5.7 is designed based on this concept, of which the one-dimensional illustration is shown in Fig. 5.14a [22]. In this configuration, if we apply an actuation force in the positive y-direction, as shown in Fig. 5.14b the compliant mechanisms will pull the table down in negative y-direction, making the dynamic effects of the actuator and the table balanced. When under shock conditions, as shown in Fig. 5.14c, an acceleration will be exerted on the whole structure. If $m_{\text{actuator}} \times x_1 = m_{\text{table}} \times x_2$, then the pivot, table, and actuator do not move relative to the frame. With this, the operation precisions of the nano-positioning system will not be affected by external shock and vibrations.

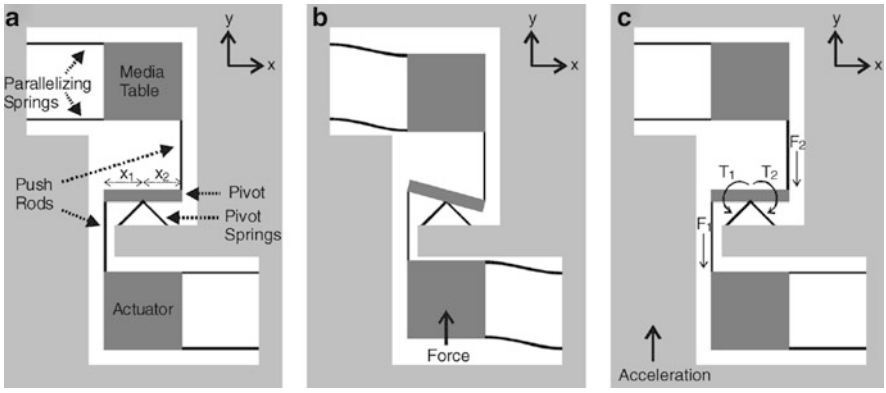


Fig. 5.14 (**a**) One-dimensional illustration of a mass balance positioner. (**b**) Rotation of pivot and displacement of the actuator and table due to the application of a force to the actuator. (**c**) Illustration of the forces and torques acting on the pivot resulting from the acceleration of the frame of the positioner [22]

5.3.2 Design Methods

The design of compliant guiding mechanisms is essential in achieving high performance for various positioning applications. However, current design of compliant guiding mechanisms stems from designer's intuitive and trial and errors. A lack of reliable and systematic design methods makes it not possible to design the compliant guiding mechanisms with optimal performance. Moreover, current designs are based on linear assumption of material properties. This is only adapted to small deformation applications and results in only a limited range of positioning motions. Thus in this section, we present two major design methods that is useful in the large deformation design and optimal design of the compliant guiding mechanisms.

5.3.2.1 Pseudo-Rigid-Body Model

Pseudo-rigid-body model is an efficient tool for analysis and design of compliant mechanisms with the assumption of large and nonlinear deflections [17]. Although there are other approaches available for large deflection problems such as elliptic integral and nonlinear finite-element methods, the formulations are cumbersome and solutions are time-consuming. The pseudo-rigid-body model, on the other hand, provides a rather simple method for large and nonlinear deflection modeling in that the rigid-link mechanism theory can be used to analyze and design the compliant mechanisms. The essential of the modeling concept is: the deflection of compliant mechanisms is modeled using rigid-body components, and the lumped springs are employed to simulate the equivalent force–deflection characteristics. Different compliant mechanisms require different pseudo-rigid-body models to accurately predict the deflection path and force–deflection relationships. In achieving this, the key issues of pseudo-rigid-body modeling are to decide where to place the pin joints and the value of the spring constants. In what follows, we present two common flexible segments that may be useful in design the compliant guiding mechanisms when large and nonlinear deflections are needed.

First we consider the small-length flexural pivots, as shown in Fig. 5.15 [17]. The beam consists of two segments: one is short and flexible, and the other is long and rigid. This beam can be treated as lumped compliant mechanism if the small segment is significantly shorter and more flexible than the large segment. In this case the pseudo-rigid-body model can be defined as shown in Fig. 5.16 [17], where the two rigid links joined at a pin joint. The pin joint is called characteristic pivot. The pivot can be put at any point along the flexible segment with no loss of accuracy because the length of the flexible segment is much small compared to that of the rigid link. For simplicity, however, the characteristic pivot is usually placed at the center of the short segment. Apparently, the accuracy of the model will be improved for the decreased length of flexible segment. The spring constant value can be found from elementary beam theory. The pseudo-rigid-model is more accurate if the bending is the dominant loading in the flexural pivot. If the transverse and axial

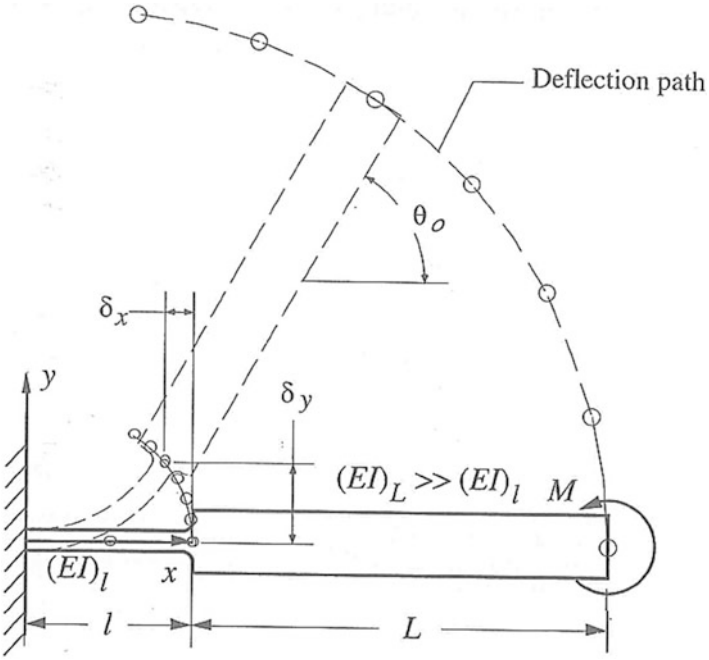


Fig. 5.15 Small-length flexural pivot [17]

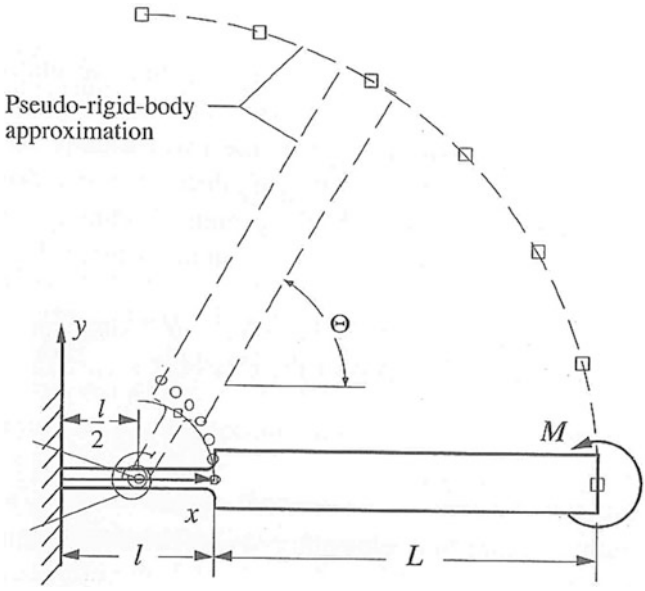


Fig. 5.16 Simple pseudo-rigid-body model of the small-length flexural pivot [17]

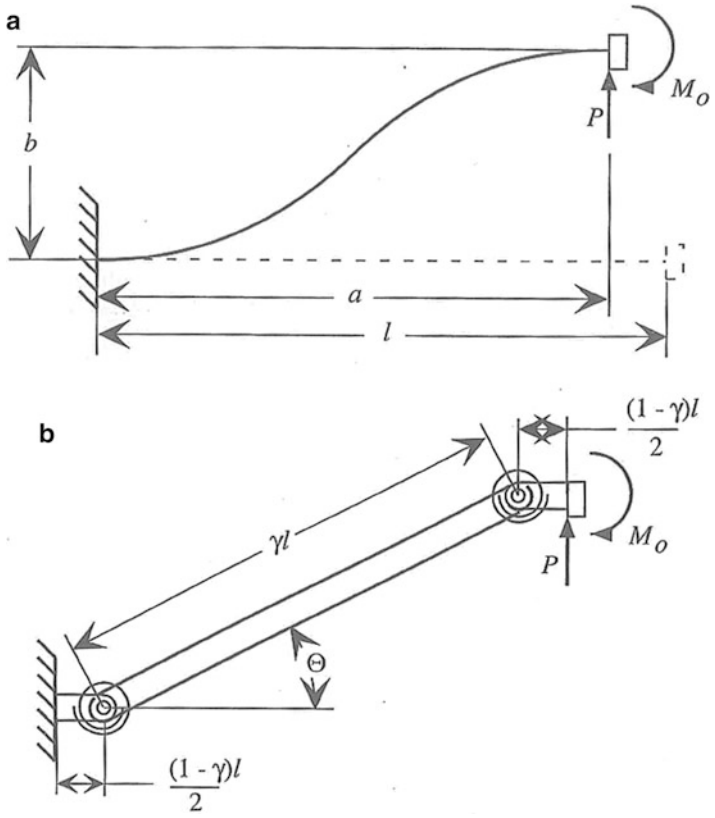


Fig. 5.17 (**a**) Fixed-guided flexible segment and (**b**) its pseudo-rigid-body model [17]

loads become significant, greater errors will be introduced using this simple pseudo-rigid-body model. Detailed descriptions of this model are presented in reference [17].

The electro-magnetic nano-positioning stages are often guided by the fixed-guided a flexible segment, as shown in Fig. 5.17a, of which one end of the beam is fixed while the other maintains at a constant angle. The associated pseudo-rigid-model is shown in Fig. 5.17b. Details of the pseudo-rigid-model can be found in [17]. It should be noted that for analysis and design of certain compliant mechanisms, the pseudo-rigid-body model is not unique. Actually, all the compliant mechanisms can be thought of as a sequence of infinite number of rigid bodies linked by springs. In modeling the compliant mechanisms, one has to make a compromise between the accuracy of the pseudo-rigid-body model and complexity of analysis and design. Fewer springs are preferred in pseudo-rigid-body model if it fulfills the accuracy requirements. As a result, the pseudo-rigid-body model conducts the analysis and design of the compliant mechanisms based on the methodologies used for rigid-body mechanisms. The modeling method is not only

suited to the kinematics but also the dynamics of compliant mechanisms. The main drawbacks of the pseudo-rigid-body model is that it only applies to compliant structure with regular geometry such as beams with constant cross-sectional area so that the elementary beam theory can be used to determine the stiffness constant.

5.3.2.2 Optimal Synthesis with Continuum Models

As mentioned, pseudo-rigid-body model and beam theory are suited only for structure that has regular geometry. However, for compliant mechanisms design, the structural geometry is not necessarily to be regular. Analysis, modeling, and design of compliant mechanisms with irregular geometry can be accomplished by using continuum model instead [23–25]. Based on continuum solid mechanics, structural optimization techniques can be used for compliant mechanisms design. By specifying the desired force–deflection behavior, materials, and design domain (i.e., the region the mechanism to be designed must fit), the structural optimization method generates the optimal solution automatically. The optimal design can be based on many design criteria such as maximum stiffness, minimum weight, suitable flexibility, and natural frequency, etc. For irregular geometries, the elastic deflection behavior is governed by partial differential equations, of which closed-form solutions are not possible. These equations are solved using numerical methods such as finite-element method, finite difference method, boundary element method, etc. The detailed introduction of this method can be found in [17].

5.4 Kinematics and Dynamics

Previous sections introduce the design issues of the two main components of the electro-magnetic nano-positioning systems, i.e., the electro-magnetic actuators and the guiding mechanisms. In the analysis and design of a whole electro-magnetic nano-positioning system, however, the primary issue is the kinematic analysis, for example, deciding how many actuators with single- or multi-axes are required in achieving stage motions with certain DOFs, and how the actuators should be arranged. In view of control, it is important to understand that how these actuators should be controlled in achieving desired motion. In addition to kinematics, some dynamics issues should also be taken into account.

5.4.1 Degrees of Freedom

The DOFs of the positioning system is essential in the study of kinematics and dynamics. In general, stage motion with several DOFs requires at least an equal number of properly arranged independent actuation inputs [26]. Unlike traditional

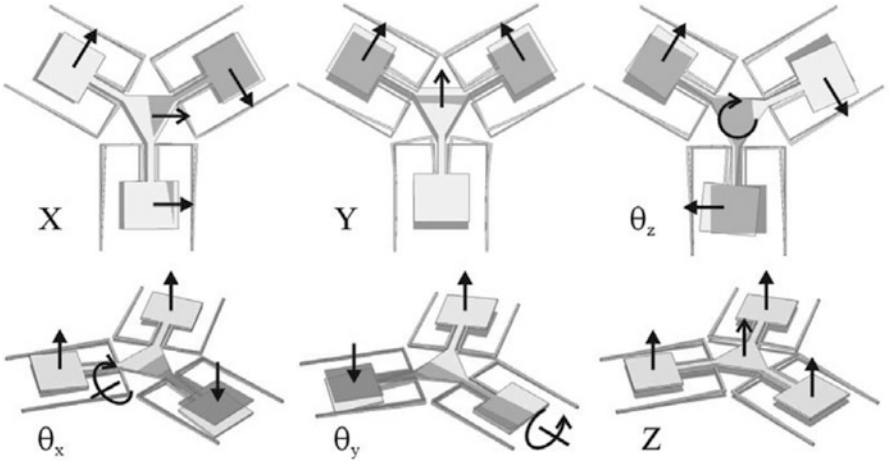


Fig. 5.18 Working principle of a six-axial positioning stage [27]

mechanisms that is capable of achieving rotations by directly using the rotation actuators. The rotation motion of the electro-magnetic nano-positioning stage is only accomplished based on the couple, because the electro-magnetic input is always a force rather than a torque. Thus a combination and arrangement of these actuation forces becomes the primary design consideration. Figure 5.18 [27] shows a typical combination and arrangement of actuations for six-axial stage applications. The upper three in-plane actuation forces are combined to generate in-plane motions, i.e., from left to right, *x*- and *y*-axial rectilinear motions, and rotation with respect to *z*-axis. The lower three vertical actuation forces are combined to generate out-of-plane motions, i.e., from left to right, rotations about *x* and *y* axes, and translation along *z* axis. These actuation forces are combined in symmetric manner to create independent motions of the stage, i.e., three independent axial motions and three independent rotations about each axis. Thus the one-to-one mapping between actuation forces and stage motions can be built. Similar layouts of the actuation forces are presented with details in [26].

5.4.2 Kinematics

The kinematics describes the relationship between the stage motions and actuation input motions. Different from conventional mechanisms kinematics such as kinematics of robotics that the kinematics are governed by geometric relationships caused by the rigid links, the kinematics of electro-magnetic nano-positioning systems are difficult to define because the motions are mainly governed by the stiffness of the compliant guiding mechanisms. In other words we need to consider the relationship between input actuation forces and resultant stage motions rather

than only the relationship between input and output motions. In current research, assumptions of linear deformations of compliant guiding mechanisms, small stage displacements, and constant actuation forces over the range of motion are made. Euler's beam theories and Castigliano's second theorem can be used to derive the stiffness of compliant guiding mechanisms that has regular geometries. For irregular geometries, closed-form solutions may be challenging. Numerical solutions such as finite-element analysis may be used to determine the stiffness. Because of these assumptions, matrices are employed for building the input–output linear kinematic mappings.

As in the six-axis positioning stage layout, here, we employ the stiffness matrix K of the compliant guiding mechanism which relates the actuation force $\mathbf{F}$ on the stage to the stage displacements $\mathbf{x}$, as in (5.6) [8]. The actuation force on the stage is related to the actuation force acting on the paddle through force transformation matrix T, and the actuation force on the paddle is related to the input vector $\mathbf{i}$ (voltage, current, etc.) through actuation matrix, K_a. K_a is derived from the electromagnetic analysis. With this, the actuation force on the stage is given by (5.7) [8].

$$\mathbf{F} = K\mathbf{x} \tag{5.6}$$

$$\mathbf{F} = TK_a\mathbf{i} \tag{5.7}$$

Equating (5.6) and (5.7), we derive the system input–output relationships, as in (5.8) [8]. For the objective stage motions, one can calculate the required input vectors using (5.8) if these matrices are invertible. Because the stage center of mass is on the plane of the horizontal actuations, the horizontal and vertical motions are decoupled.

$$\mathbf{i} = (TK_a)^{-1}K\mathbf{x} \tag{5.8}$$

5.4.3 Dynamics

With the assumption of linear system and lumped stage mass and damping, the dynamics of the system is given by [8]

$$M\ddot{\mathbf{x}} + B\dot{\mathbf{x}} + K\mathbf{x} = TK_a\mathbf{i} \tag{5.9}$$

where M, B, and K are the mass, damping, and stiffness matrices of the system. Coupling between the in-plane and out-of-plane motions generates unwanted erroneous motions and add complexity of control design. It can be decoupled by placing the stage center of mass on the plane of the horizontal actuations. However, as will be presented in later section, in certain applications, the coupling effect cannot be eliminated in this way, and it could only be minimized via control

methods. For most of the nano-positioning systems that are not lumped mass and guided by compliant mechanisms, it is difficult to establish closed-form dynamics equations. However, the dynamics of the system can be identified using simulations and experiments.

5.5 Control

The control of electro-magnetic nano-positioning systems varies with respect to different applications. Thus in this section, we present the control schemes on the typical application of electro-magnetic nano-positioning application, the probe-data storage device, as in Figs. 5.1 and 5.7. In this system [28], a MEMS-based micro-scanner with *x*/*y*-directional motions is used to position the storage medium relative to the array of probes so that data is written/read/erased. Each actuator consists of a pair of magnets and a solenoid in the middle. Upon application of a current, one magnet is attracted and the other is repelled by the solenoid, forcing the shuttle to move in the same direction. If the direction of current is altered, the shuttle will be pushed in opposite direction. Before introducing the control schemes, we first present sensors that are used to provide the positional information of the micro-scanner.

5.5.1 Position Sensors

5.5.1.1 Thermal Position Sensors

The working principle of the thermal positional sensors is depicted in Fig. 5.19 [29]. The devices have the form of "U"-shaped cantilever, which consists of two supporting legs and one resistive heater in-between. The devices are made from single-crystal silicon using standard bulk micromachining technology. The resistive heater is made from moderately doped silicon, whereas the legs are made from the highly doped silicon. When applied a voltage on the legs, current flows through the heater resulting in temperature increases of the device. To ensure a uniform temperature increases along the resistive heater, the legs are usually fabricated with equal length.

Figure 5.19c, d illustrates the working principle of the thermal position sensor. To use the device as a position sensor, it should be positioned right above the surface of the object of interest and the surface of the sensor should be parallel to the object surface. Since the larger the overlapped area between the sensor and the underlying object leads to more efficiency in cooling, the temperature of the sensor should be lower than that of the less overlapped case. On the other hand, the resistance of the sensor increases with an increased temperature and decreases with decreased temperature. This variation in resistance can be monitored by looking at the current

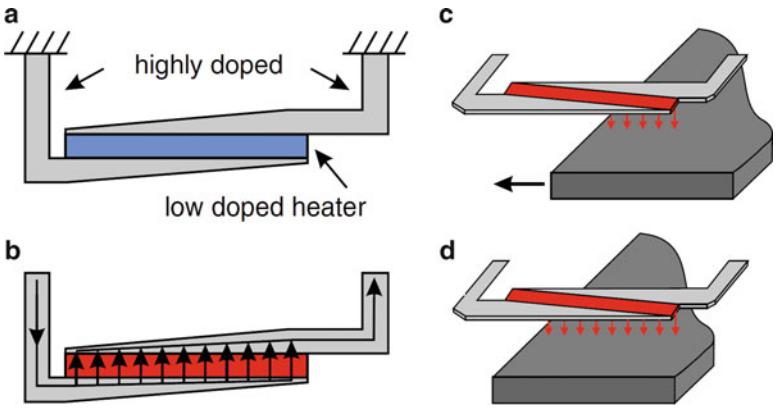


Fig. 5.19 (**a**) Schematic diagram of the thermal position sensor. (**b**) Application of a voltage on the legs results in current flow through the resistive heater. (**c**) The current causes the heat generation of the resistive heater, which is conducted to the underlying surface through the air layer. (**d**) As the underlying surface is moved towards the left, more heat is transferred to the surfaces, which results in temperature decreases of the resistive heater and its electrical resistance accordingly [29]

flowing through the beams. When the sensor is applied by a constant voltage, the change in overlapped area can be thus indirectly detected by measuring the current passing through the sensor. To obtain the position information, the long axis of the resistive heater should be aligned parallel to the direction of motion. Hence, the position can be sensed by monitoring the current flowing through the sensor. Usually for both x and y axial position sensing, two sensors are combined and arranged in orthogonal directions independently. The thermal position sensor of an electro-magnetic nano scanner is shown in Fig. 5.7b.

The single thermal sensor is capable of achieving a satisfactory resolution. However, the sensing results are affected by fluctuation in ambient temperatures and ageing effects owing to dopant diffusion or oxidation. These effects can be minimized by employing a pair of same thermal sensor and make a differential measurement. As shown in Fig. 5.20a, a pair of sensor positioned with an equal overlapped area will give rise to approximately equal cooling of the resistive heaters and therefore equal temperatures and resistances of the devices. As the underlying object displace towards the left, the difference in overlapped areas leads to difference in temperature and resistance of the resistive heater accordingly. In this arrangement, the fluctuation in ambient temperature and ageing effects for both sensors are approximately equal and thus the difference between the resistances will not be affected.

5.5.1.2 Medium-Derived PES

The thermal position sensors provide universal position information with nanometer resolutions. The probe-storage device, on the other hand, provides an alternative

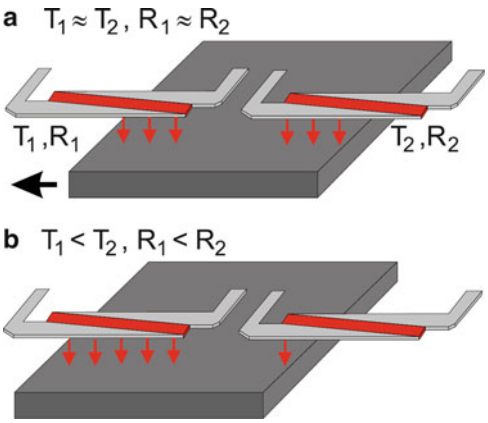


Fig. 5.20 Differential measurement set-up using a pair of matched thermal sensors. (**a**) The test object is positioned between the two thermal sensors leading to equal cooling, and results in approximately equal resistances of the two sensors. (**b**) Because the object is moved towards the left, the cooling of the left sensor increases and the right sensor decreases, and this give rise to a decrease in resistance of the sensor on the left (R_1) and increase of the sensor on the right (R_2) [29]

position sensing technique that is convenient in measuring the deviation of the read/write microcantilever tip from the track center. In addition to data fields, certain storage fields (called servo field) and the corresponding cantilever tips are reserved for generation of position signal, as in Fig. 5.21a. The servo field is shown in Fig. 5.21b, where the circle represents the indentations inscribed prior to the regular operation of the device. The *x*-direction distance of indentations in all bursts is equal to the bit pitch (BP). The cross-track distance between indentations center line of the same burst is equal to the track pitch (TP), whereas the distance between burst centers of A and B (or C and D) is TP/2 [30].

These prewritten indentations provide medium-derived position-error signal (PES), which measures the position deviation in cross-track direction. In regular operations, the track centerline of the data fields is aligned with that of the burst C. Bursts A and B are combined to create the in-phase signal (I-signal), and burst C and D are used to generate quadrature signal (Q-signal). Moreover, the PES is obtained over discrete time step which is needed by the scanner to move a BP in the *x*-direction. Thus the accurately data writing or reading operations along *x*-direction is ensured by medium-derived PES. Compared with the thermal position sensors that are capable of offering absolute position information, the medium-derived PES provides only the relative position deviations in cross-track directions. In other words, the thermal position sensors have a larger range of operation, whereas the medium-derived PES confines to a limited range of sensing. The two sensors provide similar resolutions.

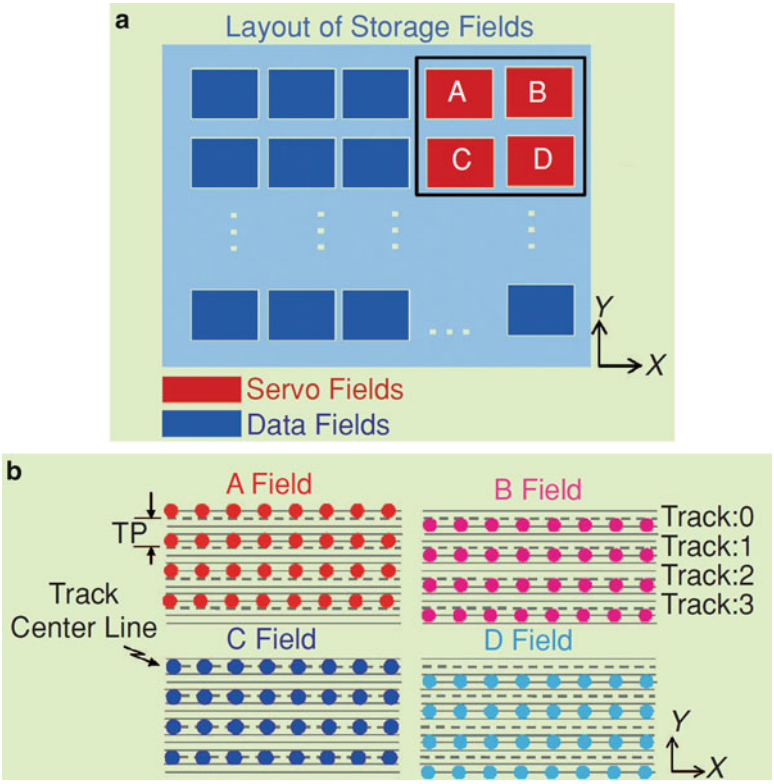


Fig. 5.21 Concept of the medium-derived PES [30]

5.5.1.3 Noise Characteristics

In the absence of external disturbances, positioning accuracy is markedly affected by the noise characteristics of the sensors. For the thermal position sensor, the predominant noise sources are Johnson noise and $1/f$-noise of the silicon resister, whereas the predominant noise sources of the medium-derived PES are medium and channel noise [30]. Moreover, the thermal position sensor suffers from drift and sensitivity variations. The noise spectral characteristics of both the thermal position sensor and medium-derived PES are shown in Fig. 5.22 [30]. As can be seen from Fig. 5.22, there is a significant low-frequency component of the thermal position sensor. This will result in the measurement particularly noisy at low frequencies. In contrast, the medium-derived PES shows good low-frequency fidelity. It should be noted that beyond a certain frequency, performance of the thermal position sensor is better than that of the medium-derived PES.

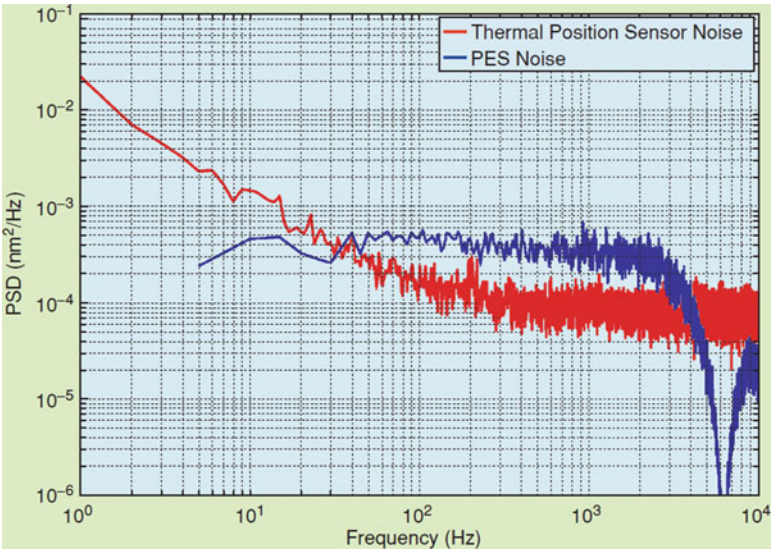


Fig. 5.22 The characteristics of the thermal position sensors and the medium-derived position-error signal (PES) [30]

5.5.2 Control Schemes

The control schemes are based on actual requirements in positioning. As with the PDS devices, the control system has mainly two functions. First, the scan table is to be located to a target track from an arbitrary initial position. This procedure is known as seek-and-settle. Second, from the target position, the medium scanner should be moved with constant velocity along the length of the track along which the writing/reading are operated. The probes should be maintained on the center of the target track while operation. This is achieved by the track-follow procedure.

5.5.2.1 Cross-Coupling

In this planar configuration of the electro-magnetic nano-positioning system, *x*- and *y*-directional motion is achieved independently by their respective actuators. However, there are cross-coupling components between the axes in the system owing to the compliant guiding mechanisms. As shown in Fig. 5.23 [31], the cross-coupling is nonlinear and position-dependent. In design of control system, the cross-coupling is usually seen as the disturbance that has to be eliminated by the control.

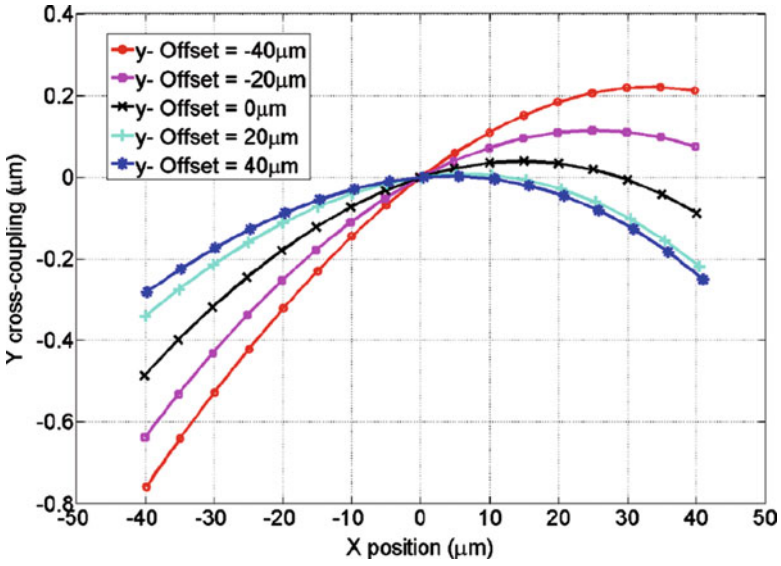


Fig. 5.23 Cross-coupling of *x*- and *y*-axes [31]

5.5.2.2 Seek-and-Settle

The seek-and-settle procedure is essential in data access time, and thus minimizing the duration in this procedure becomes an important control objective. Since medium-derived PES is only capable of providing deviation information in cross-track direction, in this seek-and-settle control procedure the thermal position sensors are thus employed to form a close-loop control scheme. Detailed control architecture can be found in [32].

5.5.2.3 Track-Follow

Different from the seek-and-settle control that the control of *x*- and *y*-directional motions of the scanner is equivalent, the track-follow control exhibits different control requirements in the two axes. As mentioned earlier, the *x*-direction scan should be kept at constant velocity to avoid timing jitter in the readback signal. Fortunately, low-frequency of positioning error is tolerable thanks to the timing-recovery circuits. This makes using the thermal position sensor acceptable despite its unreliability in measurement of errors at low frequencies. For *y*-direction control, the main objective is to maintain the probes to be at the track centerline while scanning in *x*-direction. In achieving this, both the low and high frequencies have to be minimized. The thermal position sensors, therefore, cannot be fully trusted because of the low-frequency noise, drift, and sensitivity variations. However, it is advantageous to use the thermal position sensor due to its virtues of large range and

great precision at high frequencies. To avoid the low-frequency noise induced by using the thermal position sensors, the medium-derived PES can be utilized instead.

The two-sensor-based control strategy [32], also known as the frequency-separated control [15, 30], is usually employed for the track-follow procedure. This control scheme is to achieve best error measurement by using thermal position sensor in high frequencies and the medium-derived PES in low frequencies. Moreover, the medium-derived PES can be exploited to estimate drift and make correction of using thermal position sensor at low frequencies.

5.6 Fabrication

Fabrication of macro-scale electro-magnetic nano-positioners is accomplished based on traditional approaches, and in this section we only focus the fabrication of meso- and micro-scale positioning systems. The fabrication of these positioning systems can be divided into two main parts: fabrication of the micro coil and the compliant guiding mechanisms. The compliant guiding mechanisms or flexure-supported platform is usually fabricated using a standard silicon-on-insulator MEMS (SOI-MEMS) micromachining technology. Figure 5.24 [33] shows the fabrication process of a flexure-supported platform.

In fabrication of the flexure-supported platform, the most important consideration is the thickness of the wafer. Since for planar scanning applications, out-of-plane stiffness should be ensured. In this regard, the wafer should be fabricated thicker. However, thicker wafer increases the etching time as a tradeoff. For out-of-plane applications, the structure should be fabricated more but not too flexible to both ensure out-of-plane motion and eliminate the gravity-induced errors.

Fabrication of micro coil is achieved based on two approaches, as in Fig. 5.25 [10]. One approach uses the patterned micro-mold to fabricate Cu coil with high aspect ratio, shown in Fig. 5.25a. There are various mold materials available such as SU-8, polyimide, photoresist, and silicon, etc. Fabrication procedure consists of mainly three steps: mold patterning, electroplating, and removal of mold. However, fabricating a high aspect ratio mold and removing the mold after electroplating are difficult. Another approach is to fill Cu inside Si trench by electroplating and remove the overflowed Cu using CMP (chemical mechanical polishing). The mold of high aspect ratio Cu coil is easily fabricated using ICP-RIE (inductively coupled plasma reactive ion etching) with high resolution, which ensures a lower resistance coil fabrication. In addition, the flat surface thanks to the CMP makes it much easier for other layers to assemble compared with the conventional approach. The procedure can be divided into three steps: seed layer deposition, electroplating inside Si trench, and CMP.

In constructing the whole system, one approach is to assemble the two chips (one with the micro coil and the other one with the flexible-supported platform)

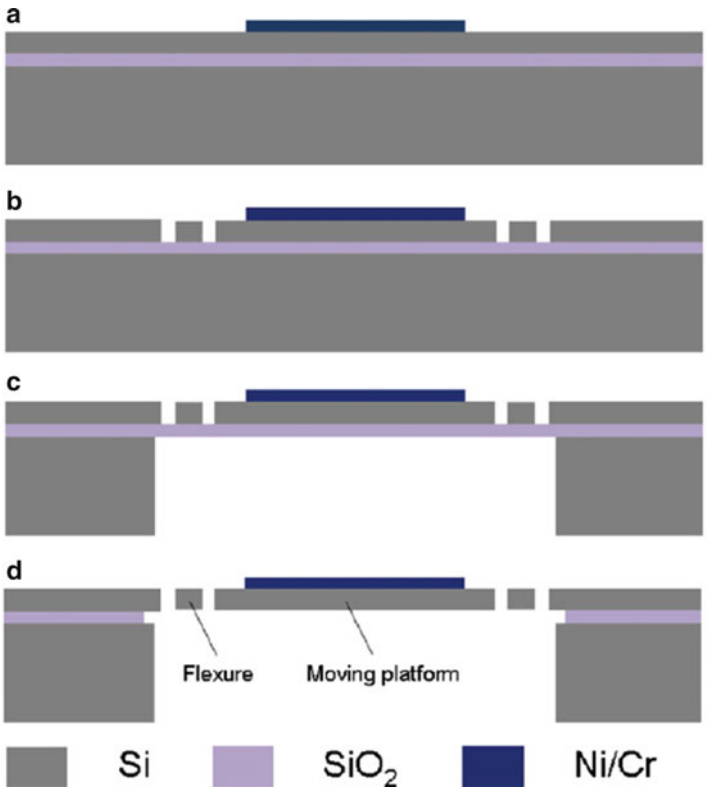


Fig. 5.24 Fabrication process for flexure-supported platform. (**a**) Ni/Cr layer are deposited using liftoff, (**b**) platform shape is etched in the device layer using DRIE, (**c**) handle wafer is etched using DRIE, and (**d**) platform is released by etching the SiO_2 layer in HF and critical point drying [33]

together with the aid of microscope. To ensure high electro-magnetic force, the distance between the coils and magnets should be minimized. Another approach is to integrate the fabrications of micro coils and flexible platform together. The choice of these two approaches depends on different applications.

5.7 Conclusions

Electro-magnetic nano-positioning technology is a relatively young field in nano technologies. This chapter presents a general introduction and guidance on electro-magnetic nano-positioning the aspects covered involve electro-magnetic actuation, guiding mechanisms, system kinematics and dynamics, control, and fabrications. These aspects focus mainly on the most common and general concepts and principles of various electro-magnetic nano-positioning systems. Therefore, potential

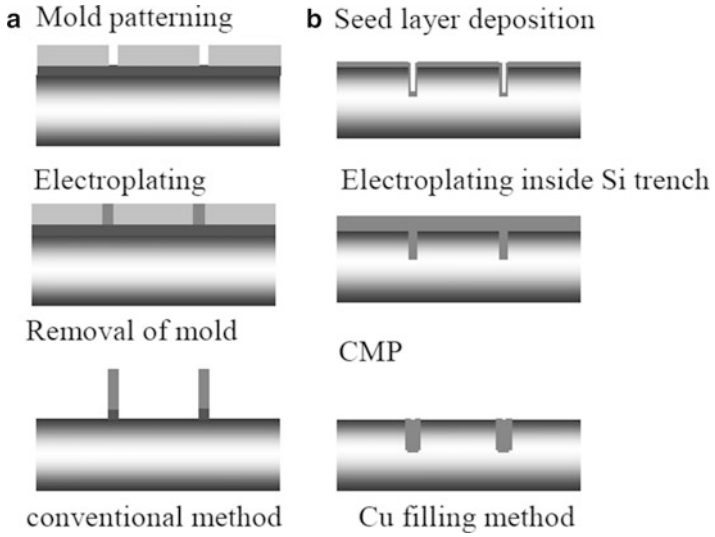


Fig. 5.25 Fabrication and process of Cu coil [10]

readers may benefit from it via an overview understanding of the nano-positioning systems. The aim is to provide researchers with guidance on developing high-performance nano-positioning systems for certain applications. In achieving this, a variety of configurations of electro-magnetic nano-positioning systems are reviewed and used as examples to illustrate the related methodologies and technologies. Critical issues for each step in developing the systems are reviewed and summarized. Not limited to current research, this chapter also presents some potential methods such as the pseudo-rigid-body model that may be helpful in the design and analysis of electro-magnetic nano-positioning systems.

References

1. D. Golda, M. Culpepper, Multi-axis electromagnetic moving-coil micro-actuator, in *ASME 2006 International Mechanical Engineering Congress and Exposition*, no. Imece2006-15191 (2006), pp. 57–63
2. G. Haugstad et al., *Scanning Probe Microscopy* (Wiley, 2013)
3. E. Eleftheriou et al., Millipede—a MEMS-based scanning-probe data-storage system. IEEE Trans. Magn. **39**(2), 938–945 (2003)
4. S. Naberhuis, Probe-based recording technology. J. Magn. Magn. Mater. **249**, 447–451 (2002)
5. A. Malshe, K. Virwani, K. Rajurkar, D. Deshpande, Investigation of nanoscale electro machining (nano-EM) in dielectric oil. CIRP Ann. Manuf. Technol. **54**(1), 175–178 (2005)
6. M. Khatib, MEMS-based storage devices: integration in energy-constrained mobile systems. Doctoral Dissertation, University of Twente, 2009
7. M. Khatib, E. Miller, P. Hartel, Workload-based configuration of MEMS-based storage devices for mobile systems, in *Proceedings of the 8th ACM & IEEE Conference on Embedded Software* (October 2008)

8. D. Golda, Design of high-speed, meso-scale nanopositioners driven by electromagnetic actuators. Doctoral Dissertation, MIT, February 2008
9. S. Verma, W. Kim, J. Gu, Six-axis nanopositioning device with precision magnetic levitation technology. IEEE/ASME Trans. Mechatronics **9**(2), 384–391 (2004)
10. J. Choi, H. Park, K. Kim, J. Jeom, Electromagnetic micro x-y stage for probe-based data storage. J. Semicond. Technol. Sci. **1**(1), 84–93 (2001)
11. C. Jaejoon, P. Hong sik, K. Kim, J. Jeon, Electromagnetic micro x-y stage with very thick Cu coil for probe-based mass data storage device. Proc. SPIE **4334**, 363–371 (2001)
12. M. King, B. Zhu, S. Tang, High-precision control of a maglev linear actuator with nanopositioning capability. Mob. Robots **8**(2), 520–531 (2001)
13. W. Kim, H. Maheshwari, High-precision magnetic levitation stage for photolithography, in *Proceedings of the American Control Conference* (2002), pp. 4279–4284
14. W. Kim, D. Trumpert, High-precision planar magnetic levitation. Precis. Eng. 66–77 (1998)
15. A. Pantazi et al., Control of MEMS-based scanning-probe data-storage devices. IEEE Trans. Control Syst. Technol. **15**(5), 824–841 (2007)
16. M. King, B. Zhu, S. Tang, 6-Axis electromagnetically-actuated meso-scale nanopositioner. IEEE Control Syst. Mag. 26–35 (2008)
17. L. Howell, *Compliant Mechanisms* (A Wiley-Interscience Publication, 2001)
18. Q. Xu, Design and development of a flexure-based dual-stage nanopositioning system with minimum interface behavior. IEEE Trans. Autom. Sci. Eng. **9**(3), 554–563 (2012)
19. Y. Yong, S. Moheimani, B. Kenton, K. Leang, Invited review article: high-speed flexure-guided nanopositioning: mechanical design and control issues. Rev. Sci. Instrum. (83) (2012)
20. G. Anderson, A six degree of freedom flexural positioning stage, Masters Thesis, MIT, 1999
21. M. Culpepper, G. Anderson, P. Petri, HEXFLEX: a planar mechanism for six-axis manipulation and alignment. Patent pending, Massachusetts Institute of Technology, Cambridge, 2001
22. M. Lantz et al., A vibration resistant nanopositioner for mobile parallel-probe storage applications. J. Microelectromech. Syst. **16**(1), 130–139 (2007)
23. S. Kota et al., Design of compliant mechanisms: applications to MEMS. Analog Integr. Circ. Sig. Process. **29**, 7–15 (2001)
24. S. Deepak, M. Dinesh, D. Sahu, G. Ananthasuresh, A comparative study of the formulations and benchmark problems for the topology optimization of compliant mechanisms. J. Mech. Robot. **1** (March 2008)
25. M. Bendsoe, O. Sigmund, *Topology Optimization Theory* (Springer, 2013)
26. J. Zhao, K. Zhou, Z. Feng, A theory of degrees of freedom for mechanisms. Mech. Mach. Theory **39**, 621–643 (2004)
27. C. DiBiasio, M. Culpepper, Design and characterization of a 6 degree-of-freedom meso-scale nanopositioner with integrated strain sensing, in *ASPE Proceedings, Design of Precision Machines* (2010)
28. S. Devasia, E. Eleftheriou, S. Moheimani, A survey of control issues in nanopositioning. IEEE Trans. Control Syst. Technol. **15**(5), 802–823 (2007)
29. M. Lantz, G. Binnig, M. Despont, U. Drechsler, *A Micromechanical Thermal Displacement Sensor with Nanometre Resolution*, vol 15 (Institute of Physics Publishing, 2005)
30. A. Sebastian et al., Nanopositioning for probe-based data storage. IEEE Control Syst. Mag. (2008)
31. A. Sebastian et al., Achieving subnanometer precision in a MEMS-based storage device during self-servo write process. IEEE Trans. Nanotechnol. **7**(5), 586–595 (2008)
32. A. Sebastian, A. Pantazi, Nanopositioning with multiple sensors: a case study in data storage. IEEE Trans. Control Syst. Technol. **20**(2), 382–394 (2012)
33. Y. Choi et al., A high-bandwidth electromagnetic MEMS motion stage for scanning applications. J. Micromech. Microeng. **22** (2012)

Chapter 6
Modeling of Piezoelectric-Actuated Nanopositioning Stages Involving with the Hysteresis

Guo-Ying Gu, Li-Min Zhu, Chun-Yi Su, and Han Ding

Abstract In this chapter, a general skeleton on modeling of piezoelectric-actuated nanopositioning stages (PANSs) involving with the hysteresis is presented. Toward this framework, a general model, including nonlinear electric behavior, voltage-charge hysteresis, piezoelectric effect, and frequency response of the stage, is developed to completely characterize dynamic behaviors of the PANS with both electrical and mechanical components. For description of the hysteresis nonlinearity, we firstly give a brief survey of the available hysteresis models in the literature. Then, in order to validate the general model, the developed general model is divided into two modules in terms of a three-order linear dynamic plant preceded by an input hysteresis nonlinearity that is described by a modified Prandtl–Ishlinskii (MPI) hysteresis model as an illustration. Subsequently, parameters of the linear part and nonlinear part of the general model are identified, respectively. For the linear part, an axiomatic-design-theory-based approach is adopted to identify the linear plant. For the nonlinear part, a particle swarm optimization algorithm is utilized for identification of the MPI model on basis of the identified parameters of the linear part in the first step. Finally, the experimental and model simulation results of a commercial PANS clearly verify the effectiveness of the proposed general model and the corresponding identification method.

6.1 Introduction

Along with the rapid development of nanoscience and nanotechnology, piezoelectric-actuated nanopositioning stages (PANSs) are becoming more and more significant owing to their nanometer or sub-nanometer displacement

G.-Y. Gu (✉) • L.-M. Zhu • H. Ding
State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: guguoying@sjtu.edu.cn; zhulm@sjtu.edu.cn; hding@sjtu.edu.cn

C.-Y. Su
Concordia University, 1455 de Maisonneuve Blvd. W., Montreal, Quebec, Canada, H3G 1M8
e-mail: cysu@alcor.concordia.ca

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_6

resolution in many precision manufacturing equipments [13, 31, 65, 75], such as micro/nano-manipulations, scanning probe microscopes, and ultra-precision machine tools. A PANS system is generally composed of a flexure-hinge-based mechanism, a piezoelectric actuator, a piezoelectric drive amplifier, and a displacement sensor. Such a mechatronic system has attracted significant attentions in the literature on its system modeling, control design, and experimental applications due to the inherent complex hysteresis nonlinearity exhibited in the piezoelectric actuators [25, 26, 29, 32, 33, 52, 76, 88].

Generally speaking, the modeling of the PANS is essential to account for dynamic effects due to frequency response of the stage, voltage-charge hysteresis, and nonlinear electric behavior, which is also related and important for control design. However, in previous reported works [1, 19, 20, 22, 67], the attentions focused on different aspects with different assumptions. For example, in [1, 20, 22, 67], the developed models mainly focused on the piezoelectric actuator itself, while the electric behavior of the piezoelectric drive amplifier was simply ignored. On the other hand, although Gao [19] proposed a linear modeling approach for the PANS including the behavior of the piezoelectric drive amplifier, the hysteresis effect of the piezoelectric actuator was totally missed.

It is obvious from the above discussions, a general skeleton on modeling of the PANS is still missing. Based on the previous accumulated results, we propose a general model for the united descriptions of the PANS with both electrical and mechanical components. The proposed model reflects dynamic effects of the PANS including the nonlinear electric behavior, voltage-charge hysteresis, piezoelectric effect, and frequency response of the stage, so that any reported model in the literature can be deduced from it, depending on the different applications. To illustrate the validity of the general model, a modified Prandtl–Ishlinskii (MPI) hysteresis model is utilized as one of the hysteresis models to describe the hysteresis effect in the stage. Finally, identification and experimental results are conducted to confirm such a model.

The remainder of this chapter is organized as follows. Section 6.2 starts with description of the PANS. In Sect. 6.3, the proposed general model is detailed and the available hysteresis models are surveyed in Sect. 6.4. In Sect. 6.5, the identified method and model validation with experimental results are presented, and Sect. 6.6 concludes this chapter.

6.2 Piezoelectric-Actuated Nanopositioning Stage

6.2.1 System Description

Figure 6.1 shows an architecture of a PANS system. Therein, (1) the flexure-hinge-guided mechanism is usually employed to provide motion by the elastic deformations, which has the advantages of the monolithic structure and no sliding

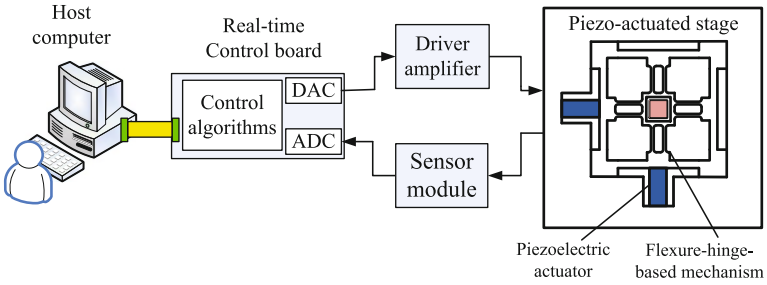


Fig. 6.1 Architecture of a PANS system

parts, thereby avoiding undesired nonlinear effects such as backlash and friction [48]; (2) the piezoelectric actuator is applied to realize the actuation by generating force on the mechanism due to its excellent advantages of the large output force, high displacement resolution, and fast response time [13, 31]; (3) the piezoelectric driver amplifier (PDA) used to supply power for the piezoelectric ceramic actuator (PCA) by either voltage control or charge control [59]; (4) a displacement sensor [16], such as the inductive, piezoresistive, capacitive, or optical sensors, with high resolution can be utilized to measure the real-time motion displacement of the stage and the sensor module is used to convert the sensor's signals to low-voltage signals for the control systems; (5) the control algorithms are implemented in the real-time control board to produce control commands for the piezoelectric actuator to achieve the nanopositioning motion of the flexure-hinge-guided mechanism.

6.2.2 Working Principle

In general, the key enabling element of a PANS is the piezoelectric actuator, which can manage small displacements in the range of sub-nanometers to several hundreds micrometers. Therefore, the main advantages of the PANSs are due to the utilization of piezoelectric actuators.

The piezoelectric actuator, made of piezoelectric materials, is a specific type of smart material-based devices that can directly convert an electrical signal into a physical displacement [31, 41, 62]. In the early time, the most applied piezoelectric materials were piezoelectric crystals such as Quartz, tourmaline, and Rochelle salt. The discovery of piezoelectric ceramic materials, such as barium titanate and lead zirconate titanate (PZT) brought a major breakthrough owing to their high Curie temperature and outstanding piezoelectric effect [41]. Compared to conventional actuators, piezoelectric actuators can be made smaller and lighter in a variety of forms, ranging from rectangular patches, thin disks, and tubes. In applications, piezoelectric actuators are generally fabricated with three shapes [62], i.e., piezoelectric bender actuators, piezoelectric tube actuators, and piezoelectric stack actuators.

Piezoelectric actuators work on the principle of piezoelectricity, which is an electromechanical coupling phenomenon between electrical properties and mechanical properties of piezoelectric materials. Mechanical stresses in the piezoelectric materials produce measurable electric charge, which is referred to the direct piezoelectric effect. Vice versa, mechanical strains are generated in response to an applied electric field and this is called the inverse piezoelectric effect.

The most famous description of the piezoelectricity was published by a standards committee of the IEEE Ultrasonics, Ferroelectrics, and Frequency Control Society with the linearized constitutive equations as follows [13, 31, 41, 58]

$$\begin{aligned} S_i &= s_{ij}^E T_j + d_{ki} E_k \\ D_m &= d_{mj} T_j + \varepsilon_{mk}^T E_k, \end{aligned} \tag{6.1}$$

where the subscript indexes $i, j = 1, 2, \ldots, 6$ and $m, k = 1, 2, 3$ refer to different directions within the Cartesian coordinate system, S and T represent the strain tensor and stress tensor, respectively, E and D are the electric field vector and electric displacement vector, respectively, s^E is the elastic compliance matrix when subjected to a constant electrical field, d is the matrix of piezoelectric material constants, ε^T is the dielectric constant matrix under condition of constant stress. The first subscript index of the d-constant gives the "electrical" direction (i.e., field or dielectric displacement) and the second one refers to the direction of mechanical strain or stress.

The equations in (6.1) essentially state that the material stain and electrical displacement are linearly affected by the subjected mechanical stress and electrical field. However, these equations fail to explicitly describe the hysteresis nonlinearity in piezoelectric actuators and the electromechanical dynamics that are observed in PANSs. Therefore, lots of electromechanical models are reported in the literature to describe the nonlinear dynamic behaviors of PANSs, which will be discussed in the following.

6.3 Comprehensive Modeling Approach of PANSs

Comprehensive dynamic modeling of PANSs refers to develop an electromechanical modeling method by completely representing the different relations (as shown in Fig. 6.2) involved in the PANS.

This method was pioneered by Goldfarb et al. [22] and Adriaens et al. [1], which was developed for the PCA. Moreover, Gao et al. [19] developed a linear modeling approach for the PANS, but the hysteresis effect of the PCA was ignored.

Fig. 6.2 Schematic representation of different relations in PANSs

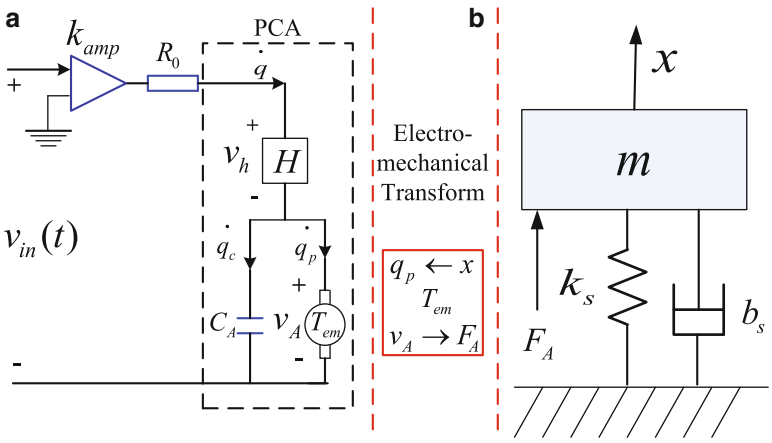


Fig. 6.3 A general model of the PANS: (**a**) electrical aspect; (**b**) mechanical aspect

Following the accumulated results developed for the PCA in [1, 19, 22], we develop a general schematic model of the PANS, which can be represented by Fig. 6.3. In the following development, we shall use this structure as a base to completely represent the characteristics of the PANS, assembling to the traditional DC motor, which will be described from both electrical and mechanical aspects.

6.3.1 Electrical Modeling

From the electrical aspect, the PANS system can be modeled as an equivalent circuit as shown in Fig. 6.3a, which is composed of a circuit of the PDA and a circuit of the PCA. From Fig. 6.3a, the PDA is modeled as a fixed gain k_{amp} plus an equivalent internal resistance R_0 of the driving circuit, v_h is the generated voltage due to the hysteresis effect $H(q)$, T_{em} represents the piezo effect, which is an electromechanical transducer with transformer ratio, and C_A represents the sum of the capacitances of the total piezoelectric ceramics, which is electrically in parallel with the transformer. The total charge in the PCA is q, and the resulting current flowing through the circuit is $\dot{q}$. The charge stored in the linear capacitance C_A is represented as q_c. The charge q_p is the transduced charge from the mechanical side due to the piezoelectric effect, the voltage v_A is the transduced voltage.

Therefore, the complete electrical equations can be expressed as [27]

$$R_0\dot{q}(t) + v_h(t) + v_A(t) = k_{\text{amp}}v_{\text{in}}(t) \tag{6.2}$$

$$v_h(t) = H(q) \tag{6.3}$$

$$q(t) = q_c(t) + q_p(t) \tag{6.4}$$

$$v_A(t) = q_c(t)/C_A \tag{6.5}$$

$$q_p(t) = T_{\text{em}} x(t), \tag{6.6}$$

where v_{in} is the control input for the PDA, x is the output displacement of the mechanical part, and k_{amp} is the fixed gain of the voltage power amplifier.

By substituting (6.3)–(6.6) into (6.2), the complete electrical equations can be further simplified as

$$R_0 C_A \dot{q}(t) + q(t) - T_{\text{em}} x(t) = C_A k_{\text{amp}} \left[v_{\text{in}}(t) - \frac{H(q)}{k_{\text{amp}}} \right]. \tag{6.7}$$

Remark 1. It is important to note it is R_0 that makes a difference in the electric part between the PANS and PCA in [1, 22]. Although the set of linear constitutive equations in [1, 22] is fine for many applications, a more realistic model requires the inclusion of R_0. Due to the existence of R_0 in the PANS, the electrical part can be expressed as a differential equation of the charge q. When $R_0 = 0$, it is reduced to the set of linear constitutive equations as discussed in [1, 22], which cannot lead to (6.7). The role of (6.7) will be further explained in the following development.

6.3.2 Electromechanical Modeling

In mechanical aspect, the following dynamic electromechanical equations can be obtained according to the piezoelectric effect and Newton's Laws of Motion

$$F_A = T_{\text{em}} v_A(t) \tag{6.8}$$

$$m\ddot{x}(t) + b_s \dot{x}(t) + k_s x(t) = F_A, \tag{6.9}$$

where F_A is the transduced force from the electrical side; x is the output displacement of the mechanical part; m, b_s, and k_s are the mass, damping coefficient, and stiffness of the moving mechanism, respectively.

To present $v_A(t)$ in (6.8) in terms of $q(t)$, it is obvious from (6.4)–(6.6) that

$$v_A(t) = \frac{1}{C_A} q(t) - \frac{T_{\text{em}}}{C_A} x(t). \tag{6.10}$$

The dynamic electromechanical equation (6.9) can then be re-written as

$$m\ddot{x}(t) + b_s\dot{x}(t) + \overline{k_s}x(t) = \frac{T_{\text{em}}}{C_A}q(t), \tag{6.11}$$

where $\overline{k_s} = k_s + \frac{T_{\text{em}}^2}{C_A}$.

6.3.3 A General Dynamic Model

Combining the electrical model (6.7) and the electromechanical model (6.11), a general dynamic model of the PANS is represented by the following set of equations

$$R_0C_A\dot{q}(t) + q(t) - T_{\text{em}}x(t) = C_Ak_{\text{amp}}\left[v_{\text{in}}(t) - \frac{H(q)}{k_{\text{amp}}}\right] \tag{6.12}$$

$$m\ddot{x}(t) + b_s\dot{x}(t) + \overline{k_s}x(t) = \frac{T_{\text{em}}}{C_A}q(t). \tag{6.13}$$

These two equations represent an analytical model of the PANS, which surprisingly assembles to the model of the traditional permanent magnet DC motor.

Remark 2.

(i) A complete comprehensive model described in (6.12) and (6.13) accounts for both the dynamics of the PANS and the hysteresis inherent to the PCA. Surprisingly, until now there is no such a complete description available in the literature though a lot of papers have addressed this issue [5, 9, 51, 52, 77]. Possible reason is that the most of available results focused on a piezoelectric actuator itself, not on the stage, i.e., [1, 20, 22, 67], where the dynamic models consisting of a second-order linear plant with hysteresis nonlinearities, focused on the PCA and ignored R_0. The proposed model can be reduced to this case if the electric circuit of the PDA is neglected, that is $R_0 = 0$ in (6.7). On the other hand, Gao [19] proposed a linear modeling approach including both PDA and PCA. However, the hysteresis effect of the PCA is ignored. The proposed model can also be reduced to this case without considering the hysteresis effect $H(q)$ in (6.12).

(ii) It is well known that dynamic model of a permanent magnet DC motor can be described as:

$$L\frac{i(t)}{dt} + Ri(t) + K_{\text{emf}}\frac{\theta(t)}{dt} = v_{\text{in}}(t) \tag{6.14}$$

$$J\ddot{\theta}(t) + B\dot{\theta}(t) = K_ti(t), \tag{6.15}$$

where i is the armature current, θ is the angular position, L, R, K_{emf}, and K_t are motor's inductance, resistance, back emf constant, and torque constant, J is the inertia of the rotor and the equivalent mechanical load, and B is the damping coefficient. From Eqs. (6.12)–(6.13), it is amused that the dynamic model of the PANS is so similar to the traditional DC motor, except the hysteresis nonlinearity $H(q)$. Therefore, the challenge for control of the PANS mainly lies on existence of the non-smooth hysteresis nonlinearity $H(q)$, which usually degrades the system performance in such manners as giving rise to undesirable inaccuracies or oscillations, even leading to instability in the closed-loop manner [83].

(iii) It is worth of mentioning that the charge control of PCA instead of the voltage control can eliminate the effect of hysteresis from (6.11), which is also demonstrated in [17]. This is exactly similar to directly control the armature current of a permanent magnet DC motor. However, as similar to the DC motor, the charge control has not been widely adopted due to the complicated implementation of the power driver and the cost of such techniques [13]. Therefore, this paper discusses the related modeling and control problems when the voltage control approach is used.

6.4 Hysteresis Modeling

As discussed in the previous section, it is the hysteresis inherent in the PCA that makes the modeling and control of the PANS challenging. In order to describe the hysteresis nonlinearity in (6.12), we will first give a brief survey of the available hysteresis models in this section.

6.4.1 What Is Hysteresis?

The word "hysteresis" is of the Greek origin, which means "lagging behind" or "coming behind." The Scottish physicist, Alfred Ewing, introduced this word into scientific vocabulary as follows [42]: "when there are two quantities M and N, such that cyclic variations of N cause cyclic variation of M, then if the changes of M lag behind those of N, we may say that there is hysteresis in the relation of M and N." As pointed out by Mayergoyz [57], this may be misleading and can create the impression that the looping is the essence of hysteresis. However, it must not be confused with "delay" or "phase lag," which is not a hard nonlinearity and is present in many linear systems.

Hysteresis in the PCA is a kind of non-smooth nonlinear phenomenon between the applied voltage and the output displacement of the piezoelectric actuator [30]. This nonlinearity is multivalued and nonlocal memoryless, which results in the major- and minor-loops behavior as shown in Fig. 6.4. It can be seen that, the

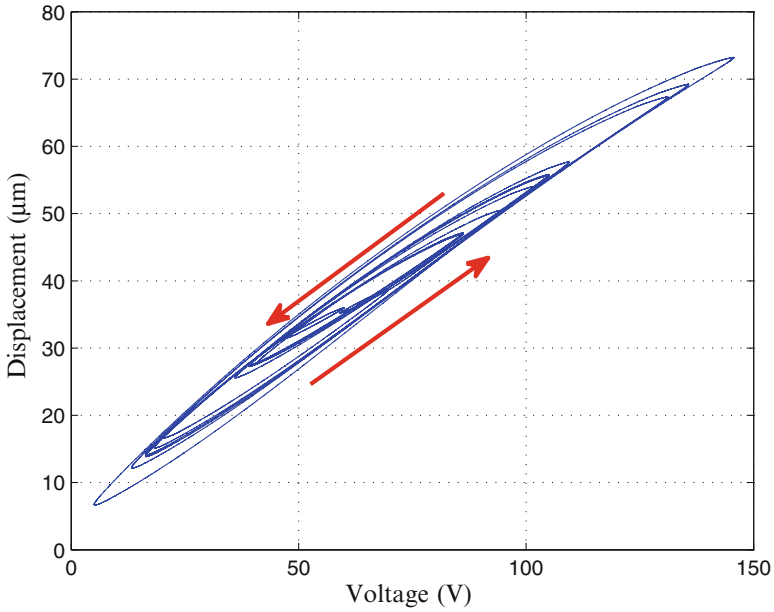


Fig. 6.4 An illustration of the hysteresis nonlinearity

displacement of the output depends upon not only the current value but also the previous dominant extrema of the input voltage. In general, the maximum error caused by the hysteresis can be as much as 15 % of the travel range. With the increase of frequencies of the input signals, the hysteresis caused errors may go beyond 35 %.

In general, for treatment of the hysteresis, the first step is to develop mathematical models that are sufficiently accurate, amenable to controller design for hysteresis compensation, and efficient enough to use in real-time applications. As a consequence, a number of hysteresis models are developed in the literature. The reader may refer to [31] for a detailed survey.

Roughly speaking, the reported hysteresis models for PANSs can be classified into two categories: physics-based models and phenomenological models. Figure 6.5 depicts the categories of these different hysteresis models.

6.4.2 Physics-Based Models

Physics-based models are built on first principles of physics effects. The Jiles–Atherton model [40, 71] is the most famous physics-based model, which was first introduced for ferromagnetic hysteresis. On the basis of the Jiles–Atherton model, Smith and Ounaies [78] proposed a domain wall model for hysteresis description

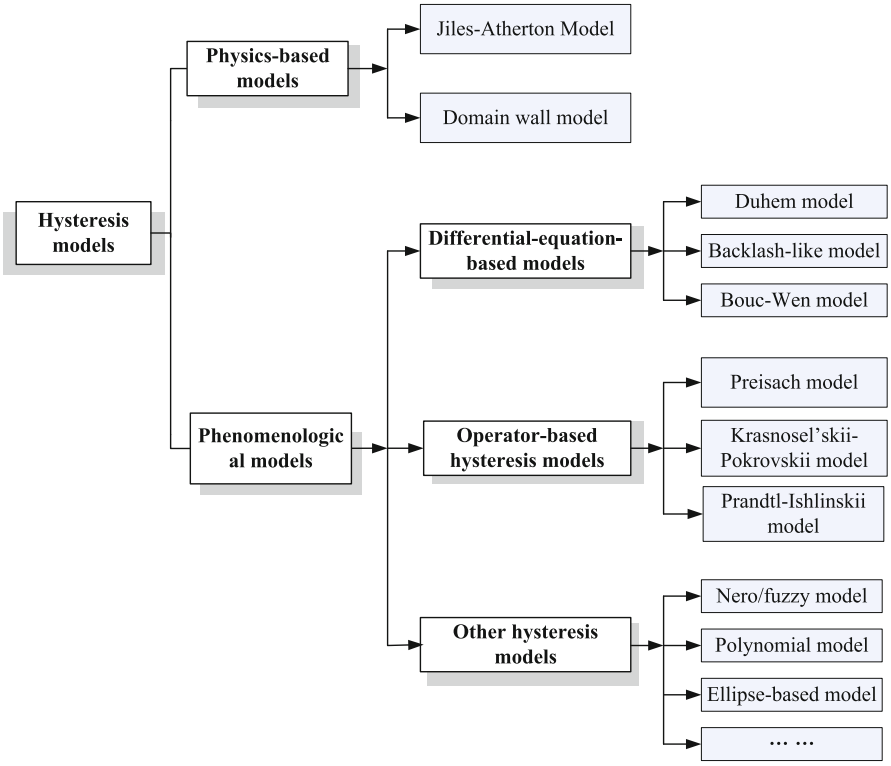


Fig. 6.5 Category of hysteresis models for PANSs

of the PCA with an ordinary differential equation formulation that was amenable to construct an inverse compensator for the linear controller design. The physics-based model is generally very complicated, and the physics-based model developed for one material may not be used for another kind of material.

6.4.3 Phenomenological Models

In contrast to physics-based models, phenomenological models are used to produce behaviors similar to those of physical models without necessarily providing physical insight into the modeling problems.

According to the differences of the mathematical structure, phenomenological models for the hysteresis description in the piezo-actuated stages can be classified as differential-equation-based hysteresis models, integral-equation-based hysteresis models (generally called operator-based hysteresis models as well) and other hysteresis models.

6.4.3.1 Differential-Equation-Based Models

Differential-equation-based models refer to a hysteresis description method in the form of differential equations [84]. As shown in Fig. 6.5, the Duhem model, Backlash-like model, and Bouc–Wen model are popular with differential equations to describe the hysteresis in PANSs.

(1) Duhem model

From 1897, the Duhem model [55] has been used for hysteresis description. Although it is originally developed to describe the magnetic hysteresis, the Duhem model is also effective to represent the hysteresis of the PCA [54, 91].

The Duhem model has the properties that every state is an equilibrium under constant inputs and the output can only change its character when the input changes direction [55, 64]. For suitable functions f_1 and f_2, the Duhem hysteresis model between the input v and output w is expressed as [55]

$$\dot{w}(t) = f_1\left(w(t), v(t)\right)\dot{v}_+(t) - f_2\left(w(t), v(t)\right)\dot{v}_-(t), \tag{6.16}$$

where

$$\dot{v}_\pm(t) = \frac{|\dot{v}(t)| \pm \dot{v}(t)}{2}. \tag{6.17}$$

In applications, the model (6.17) can be simplified to a special case [11, 55, 63]

$$\frac{dw}{dt} + a\left|\frac{dv}{dt}\right| g(v, w) = b\frac{dv}{dt}, \tag{6.18}$$

where a typical choice for g is $g(v, w) = w - c\phi(v)$.

(2) Backlash-like model

Based on the simplified Duhem model (6.18), Su et al. [81] developed a Backlash-like model, which is described as

$$\frac{dw}{dt} = \alpha\left|\frac{dv}{dt}\right|[cv - w] + B_1\frac{dv}{dt}, \tag{6.19}$$

where α, c, and B_1 are constant, satisfying $c > B_1$. Following the derivation in [81], the solution of the model (6.19) is explicitly expressed as

$$w(t) = cv(t) + d(v(t)) \tag{6.20}$$

with

$$\begin{aligned} d(v(t)) := {} & [w_0 - cv_0]e^{-\alpha(v - v_0)\mathrm{sgn}(\dot{v})} \\ & + e^{-\alpha v\,\mathrm{sgn}(\dot{v})}\int_{v_0}^{v}[B_1 - c]e^{\alpha\zeta\,\mathrm{sgn}(\dot{v})}d\zeta \end{aligned} \tag{6.21}$$

for $\dot{v}$ constant where $w_0 = w(0)$ and $v_0 = v(0)$ are the initial condition, and sgn is the well-known Sign function. The benefit for choosing the Backlash-like model is that the hysteresis nonlinearity is expressed as a linear function of the input signal plus a bounded disturbance. Thus, the conventional robust control approaches can be utilized to deal with the hysteresis without constructing the inverse of the hysteresis [27, 69, 81, 95].

(3) Bouc–Wen model

On the basis of the Duhem model (6.16), a hysteretic semi-physical model was proposed initially by Bouc early in 1971 and further generalized by Wen in 1976 [60]. Since then, the Bouc–Wen model has been known and utilized to represent the hysteresis in the form

$$\dot{w} = A\dot{v} - \beta\dot{v}|w|^n - \alpha\,|\dot{v}|\,|w|^{n-1}w, \tag{6.22}$$

where the parameters n, A, β, and α are shape parameters of the hysteresis curves. It can be seen from (6.22) that the input is $\dot{v}$ rather that v although the hysteresis is observed between v and w. The Bouc–Wen model has an interesting simplicity and is able to represent a large class of hysteresis [60]. Because of the elastic structure and material of piezoelectric ceramics, it is generally admitted that $n = 1$ to describe the hysteresis of PANSs.

From the above discussions, the differential-equation-based models can be essentially expressed as a first-order differential equation that relates the input to the output in a hysteretic way. In such a way, it can be conveniently augmented by introducing a state variable into the dynamic equations of the piezo-actuated system for controller development [27, 51, 54, 68, 91]. However, it is difficult to obtain the general solutions of differential-equation-based models. On the other hand, until now, there is no method to construct the analytical inverses of the differential-equation-based models for hysteresis compensation.

6.4.3.2 Operator-Based Hysteresis Models

Operator-based hysteresis models are kinds of models with an integral of weighed elementary hysteresis operators. Based on the differences of elementary hysteresis operators, operator-based hysteresis models for PANSs include the Preisach model, Krasnosel'skii–Pokrovkii model, and Prandtl–Ishlinskii model as shown in Fig. 6.5.

(1) Preisach model

The elementary hysteresis operator in the Preisach model is called as Relay. A geometric illustration of the Relay operator $\hat{\gamma}_{\alpha\beta}[v(t)]$ is shown in Fig. 6.6. With the Relay operator, the Preisach model is expressed as [42]

$$w(t) = P[v](t) = \int\int_{\alpha\geq\beta} \mu(\alpha,\beta)\hat{\gamma}_{\alpha,\beta}[v(t)]d\alpha d\beta \tag{6.23}$$

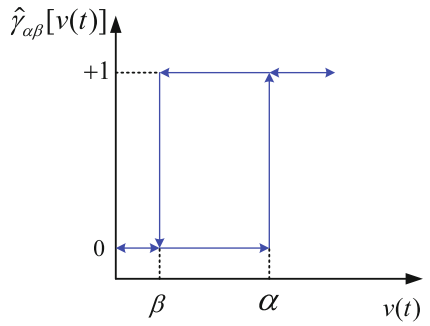


Fig. 6.6 A geometric illustration of the Relay operator

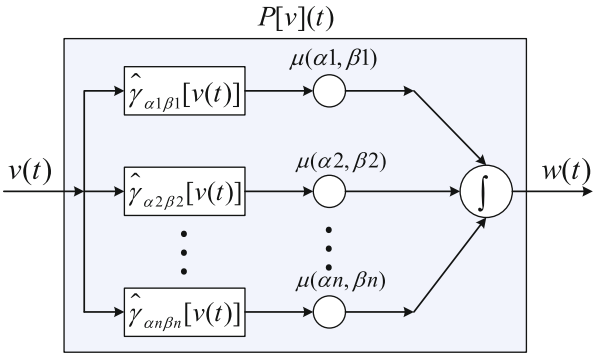


Fig. 6.7 Block diagram of the Preisach model

with

$$\hat{\gamma}_{\alpha\beta}[v(t)] = \begin{cases} 0, & v(t) \le \beta \\ 1, & v(t) \ge \alpha \\ \hat{\gamma}_{\alpha\beta}[v_0], & \forall \tau \in [t_0, t],\ v(\tau) \in (\beta,\ \alpha) \\ 0,\ \textit{if}\ v(\tau) \in (\beta,\ \alpha), & \exists t_1 \in [t_0, t),\ \text{s.t.} \\ \quad v(t_1) = \alpha \text{ and } \forall \tau \in (t_1, t]\ v(\tau) \in (\beta,\ \alpha) & \\ 1,\ \textit{if}\ v(\tau) \in (\beta,\ \alpha), & \exists t_1 \in [t_0, t),\ \text{s.t.} \\ \quad v(t_1) = \beta \text{ and } \forall \tau \in (t_1, t],\ v(\tau) \in (\beta,\ \alpha) & \end{cases}, \tag{6.24}$$

where $\mu(\alpha, \beta)$ is called as the density function, α and β are values of the switching thresholds satisfying $\alpha \ge \beta$. The simple block diagram of the Preisach model (6.23) can be approximately considered as parallel connections of a number of Relay operators as shown in Fig. 6.7.

The Preisach model was first suggested almost in the 1930s [42, 55]. A thorough discussion of the Preisach model may refer to the Mayergoyz's monograph [6]. The Preisach model can be recognized as a fundamental toolkit of hysteresis modeling for a wide range of hysteresis phenomena because of its general structure. However, due to the existence of the double integrals, identification and

implementation of the Preisach model are usually very complicated [26]. Additionally, the Preisach model is not analytically invertible. Thus, numerical methods are generally adopted to obtain approximate inversions of the model in applications [34, 35, 73, 74, 80, 85, 86].

(2) Krasnosel'skii–Pokrovkii model

The elementary hysteresis operator of the Krasnosel'skii–Pokrovkii (K–P) model is called as a KP operator [18], which is a kind of generalized plays [84]. Like the Preisach model, the K–P model is expressed using the double integrals as follows [18]

$$w(t) = \Lambda[v](t) = \int \int_{\rho_2 \geq \rho_1} \mu(\rho_2, \rho_1) k_p[v, \xi(\rho)](t) d\rho_2 d\rho_1, \tag{6.25}$$

where $\mu(\rho_2, \rho_1)$ denotes the density function, ρ_1 and ρ_2 are values of the switching thresholds satisfying $\rho_2 \geq \rho_1$. $k_p[v, \xi(\rho)](t)$ represents the KP operator with the following form

$$k_p[v, \xi(\rho)](t) = \begin{cases} \max_{\rho \in P}\{\xi_0(\rho), r(v, \rho_2)\}, & \dot{v} > 0 \\ \min_{\rho \in P}\{\xi_0(\rho), r(v, \rho_1)\}, & \dot{v} < 0 \\ k_p, & \dot{v} = 0 \end{cases} \tag{6.26}$$

with

$$r(v, x) = \begin{cases} -1, & v \leq x \\ -1 + \frac{2}{\alpha}(v - x), & x < v < x + \alpha \\ +1, & v \geq x + \alpha \end{cases}, \tag{6.27}$$

where $\rho = (\rho_1, \rho_2)$, $\alpha > 0$, k_p means the value of the KP operator does not change from its previous value, $\xi_0(\rho)$ is the initial value at location ρ satisfying $-1 \leq \xi_0(\rho) \leq 1$. Figure 6.8 shows a geometric illustration of the KP operator $k_p[v, \xi(\rho)](t)$.

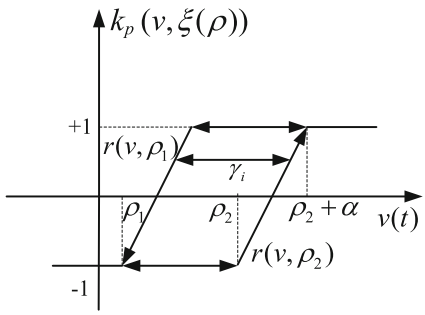


Fig. 6.8 Geometric illustrations of the KP operator

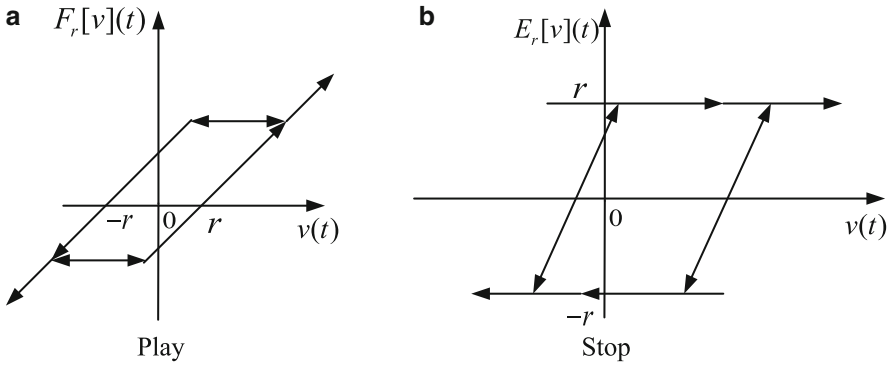


Fig. 6.9 Geometric illustrations of (**a**) Play and (**b**) Stop operators

It is pointed out that the K–P model is mainly used for hysteresis modeling and compensation in applications with shape memory alloy actuators [2, 3, 87]. Galinaities [18] investigated the properties of the K–P model and developed an approximate inverse for hysteresis compensation of a PCA. Like the Preisach model, the K–P model is also difficult to get an analytical inverse [18, 87].

(3) Prandtl–Ishlinskii model

The Prandtl–Ishlinskii (P–I) model was originally proposed as a model for plasticity–elasticity [55]. Like the Preisach model, the P–I model is defined by the elementary hysteresis operator called Play or Stop. Different from the discontinuous Relay operator with two thresholds, the Play or Stop is a continuous operator with a single threshold. Figure 6.9 shows the geometric illustrations of the Play ($F_r[v](t)$) and Stop ($E_r[v](t)$) operators. Considering the fact that $F_r[v](t) + E_r[v](t) = v(t)$, the P–I model defined by the play operator is detailed in this chapter as an illustration.

With a threshold r for any piecewise monotone input function $v(t) \in C_m[0, t_E]$, the Play-based P–I model is expressed as [7, 43]

$$w(t) = \Pi[v](t) = p_0 v(t) + \int_0^R p(r) F_r[v](t) dr \tag{6.28}$$

with

$$\begin{aligned} w(0) &= F_r[v](0) = f_r(v(0), 0) \\ w(t) &= F_r[v](t) = f_r(v(t), w(t_i)) \end{aligned} \tag{6.29}$$

for $t_i < t \le t_{i+1}, 0 \le i \le N-1$ with

$$f_r(v, w) = \max(v - r, \min(v + r, w)), \tag{6.30}$$

where $p(r)$ is a density function and p_0 is a positive constant. The density function $p(r)$ generally vanishes for large values of r, while the choice of $R = \infty$ as the upper limit of integration is widely used in the literature for the sake of convenience [30, 44].

The P–I model (6.28) is closely linked with the Preisach model (6.23) even if they seem very different at first glance. In fact, the P–I model falls into the set of Preisach model and can be considered as a subclass of the Preisach model [7, 85]. The main advantage of the P–I model over the Preisach model is that the P–I model has the analytical inverse [38, 44], thus making it more efficient for real-time applications. However, the P–I model essentially describes the symmetric hysteresis effect. It fails to characterize the hysteresis effect with the asymmetric behavior observed in the input-output relations of piezo-actuated stages. In an attempt to overcome this limitation, several variations of P–I models for describing the asymmetric hysteresis are reported. Kuhnen [45] proposed dead-zone operators cascaded with the play operator. Alternatively, Gu et al. [30] proposed a modified P–I model with a generalized input function without modifying the play operator. Instead of using the play operator, the asymmetric backlash operator [14] was adopted as the elementary operator to represent the asymmetric hysteresis behavior. As another choice, two asymmetric operators [39] were utilized for the ascending and descending branches of hysteresis loops, respectively. Janaideh et al. [37] developed two nonlinear envelope function-based operators to form a generalized P–I model for asymmetric hysteresis description. In current stage, interesting in studying P–I model for hysteresis description and compensation in piezo-actuated systems is increasing [15, 26, 28–30, 66, 70, 76].

(4) Rate-dependent model

It is pointed out that the operator-based hysteresis models are by nature developed to describe the hysteresis with the rate-independent memory effect. The rate-independent memory effect of the hysteresis model means that the output of the hysteresis model depends not on the variation rate of the input signals. In some times, the rate-independent effect is also called as the quasi-static behavior, which means the hysteresis output is not affected by the dynamics. In mathematics, the definition of the rate independence can be described by Visone [85], Krejci [43], and Brokate and Sprekels [7]: The operator $\aleph$ is defined as rate-independent if and only if the following condition holds

$$\aleph[x \circ \varphi] = \aleph[x] \circ \varphi \tag{6.31}$$

with

$$\varphi : t \in [0, t] \to \varphi(t) \in [0, t], \tag{6.32}$$

where the symbol $\circ$ is the composition operator, $\varphi(0) = 0$ and $\varphi(t) = t$. From the definition of the Relay $\hat{\gamma}_{\alpha\beta}$ (6.24) and Play $F_r[v](t)$ (6.29), Preisach and P–I hysteresis model are intrinsically rate-independent.

To overcome such limitation, Mayergoyz [56] developed a rate-dependent Preisach model by introducing the speed of output signals in the density function. Since then, a number of rate-dependent hysteresis models have employed this dynamic density function for characterizing rate dependency of the input-output relationships of piezo-actuated stages. Due to the noise problem for deriving the output signals, Yu et al. [92] proposed an alternative method by introducing the variation rate of the input signals to the density function of the Preisach model. Mrad and Hu [61] introduced a time-dependent average rate of change of the input voltage into the Preisach model to account for the rate-dependent hysteresis. With the same idea in [61], Tan et al. [82] developed a rate-dependent P–I hysteresis model. Furthermore, Janaideh et al. [36] developed a generalized rate-dependent P–I model by modifying not only the density function but also the play operator with the rate change of the input voltage. As reported, these models [36, 56, 61, 82, 92] can well predict the rate-dependent behaviors of the hysteresis in piezo-actuated stages. However, owing to the introductions of the derivatives of the input or output signals, the density function and the elementary operators become more complicated, which introduces difficulties in the model identification and real-time controller design for hysteresis compensation.

6.4.3.3 Other Hysteresis Models

Besides the differential-based and operator-based models, there are some other modeling approaches to describe the hysteresis in PANSs.

Intelligent modeling approaches such as neural networks [50, 79, 93], fuzzy systems [49], and support vector machines [89] have been applied to predict the hysteresis. Additionally, Cruz et al. [12] proposed a concept of the apparent phase shift between input and output to describe the hysteresis. The advantage of the concept is that a phaser can be conveniently used to develop hysteresis compensator. Owing to the simplicity of the mathematical description, polynomial function-based models [4, 10, 72] were developed to predict the rate-independent hysteresis. Based on the experimental phenomena of the actuator's hysteresis, Gu and Zhu [23, 24] proposed an ellipse-based hysteresis model to characterize the rate-dependent hysteresis of a PANS.

It should be noted that these hysteresis models can precisely and effectively describe the hysteresis of PANSs in particular applications, although they are not popular as the differential-based and operator-based models. On the other hand, it demonstrates that, until now, there is no a general model to completely represent the hysteresis behaviors of PANSs.

6.5 Validation of the Proposed Model

To verify the proposed general dynamic model, a PANS (PZT mode P-753.31C produced by Physik Instrumente GmbH & Co.) is used as a test bed. PZT mode P-753.31C is driven by applied control input voltage in the range of −2 to 10 V and has nominal 38 μm expansion. The PDA is a voltage power amplifier with a fixed gain of 10 to amplify the control input for driving the PCA. The unloaded resonance frequency of the mechanical response is 2.9 kHz. The capacitive gap sensor is integrated in the mechanical structure to measure the displacement of the stage. The position servo-control module (PSCM) is used as a sensor module to transfer the displacement to analogue voltage in the range of 0–10 V. The specifications of the piezoelectric positioning stage P-753.31C are listed in Table 6.1.

6.5.1 Gain of the PSCM

Before model validation, the gain K_{PSCM} of the PSCM used to transfer the displacement to analogue voltage in the range of 0–10 V needs to be fixed, which is calculated as

$$K_{\text{PSCM}} = 10/(x_{10} - x_0) \ \ (\text{V}/\mu\text{m}), \tag{6.33}$$

where x_{10} is the real stage position when the sensor monitor output is 10 V, and x_0 is the real stage position when the sensor monitor output reaches to 0 V. For this stage, $x_{10} = 38\,\mu\text{m}$ and $x_0 = 0\,\mu\text{m}$, thus one can get $K_{\text{PSCM}} = 10/38$ V/μm.

6.5.2 Identification of the Linear Dynamic Plant

It should be mentioned that, until now, there is no general approach to identify the parameters of the analytical model (6.12) and (6.13) due to existence of the

Table 6.1 Properties of the PZT mode P-753.31C

Items	Piezoelectric positioning stage	Units
Driven voltage	−2 to 10	V
Expansion displacement	0 to 38	μm
Electrical capacitance	4.6	μF
Stiffness in motion direction	16	N/μm
Unloaded resonant frequency	2.9	kHz
Stage mass	0.25	kg
PSCM output	0–10	V

hysteresis nonlinearity. As a comprise, the commonly adopted approach in the literature is to express the analytical model into a linear system preceded by a hysteresis term. Then, the linear parameters and the hysteresis model will be identified separately [21, 27, 53]. In the following development, we will follow this line.

Since the PDA operates in the voltage control mode to supply charge for the PCA to generate the piezoelectric expansion, the induced charge $q(t)$ in the circuit should be represented as a function of control input $v_{\text{in}}(t)$, that is, $q(t) = f(v_{\text{in}}(t))$. Without loss of generality, the term $v_{\text{in}}(t) - \frac{H(f(v_{\text{in}})}{k_{\text{amp}}}$ is re-defined as a new hysteresis nonlinearity $P(v_{\text{in}}(t))$

$$w(t) = P(v_{\text{in}}(t)), \tag{6.34}$$

where $P(v_{\text{in}}(t)) = v_{\text{in}}(t) - \frac{H(f(v_{\text{in}})}{k_{\text{amp}}}$. In this case, the signal from $P(v_{\text{in}}(t))$ to the output of the displacement x in (6.12) and (6.13) can be described by a third-order linear system,

$$G(s) = \frac{X(s)}{W(s)} = \frac{K}{s^3 + a_0 s^2 + a_1 s + a_2} \tag{6.35}$$

with

$$\begin{cases} a_0 = \frac{1}{R_0 C_A} + \frac{b_s}{m} \\ a_1 = \frac{b_s}{m R_0 C_A} + \frac{k_s}{m} \\ a_2 = \frac{k_s}{m R_0 C_A} \\ K = \frac{k_{\text{amp}} T_{\text{em}}}{m R_0 C_A} \end{cases}, \tag{6.36}$$

where K, $a_i, \text{i} = \{0, 1, 2\}$, are the parameters related to those parameters in (6.12) and (6.13). Thus, the general model of the PANS can be represented by a linear third-order plant (6.35) preceded by a hysteresis nonlinearity $P(v_{\text{in}}(t))$. Figure 6.10 shows the block diagram of the general model.

Remark 3. We should mention that there are two approaches to handle the hysteresis nonlinearity $H(q)$ in literature. One is by assuming $R_0 = 0$, the complete model is described by

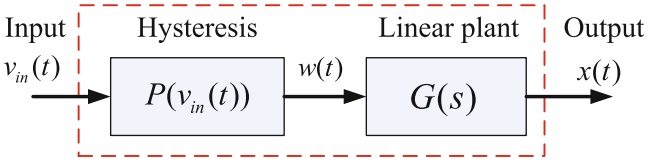


Fig. 6.10 Block diagram of the general model

$$m\ddot{x}(t) + b_s\dot{x}(t) + k_s x(t) = T_{\text{em}}[k_{\text{amp}} v_{\text{in}}(t) - H(q)]. \tag{6.37}$$

Then, the Bouc–Wen model is utilized to describe the hysteresis term $H(q)$, as in [51, 53]. The second approach is to treat the term $v_{\text{in}}(t) - \frac{H(f(v_{\text{in}})}{k_{\text{amp}}}$ as a new hysteresis nonlinearity [5, 9, 94]. For different approaches, the treatment for the controller designs is quite different because the hysteresis term in the first can be regarded as a sum component while the second is as a preceding component.

From (6.35) and (6.36), the parameters for the linear dynamic part cannot be obtained by simply using the specifications of the PANS, because the exact information on the parameters b_s, T_{em}, and R_0 is generally not available. They are usually identified based on the applied input voltage and corresponding displacement. In this work, the axiomatic-design-theory-based approach (ADTBA) [27] is utilized to identify parameters (K and $a_i, \text{i} = 0, 1, 2$) of the linear plant (6.35) using experimental data of both frequency response and transient response. For model identification of the linear dynamic part, the third-order linear dynamic model is generally expressed as $k_p{\omega_n}^2/((\tau s+1)(s^2+2\varsigma\omega_n s+{\omega_n}^2))$, where four parameters, i.e., k_p, ω, ς, and τ, should be determined. According to the ADTBA, the frequency response data is used to determine the parameters k_p and ω_n, and the transient response data is used to determine the parameters ς and τ. It is worthy of mentioning that small-amplitude input excitation signals should be applied to avoid distortion from the preceded hysteresis nonlinearity as much as possible and the higher sampling frequency (i.e., $f_s = 20\,\text{kHz}$) is set to capture the fast dynamic response of the stage when the creep behavior can be conservatively neglected. The detailed steps are described below:

(1) A small-amplitude band-limited white noise signal is used to excite the PCA. The system identification toolbox of MATLAB is adopted to obtain the frequency response of the PANS as shown in Fig. 6.11 indicated by the solid blue line. The parameters k_p and ω_n can be directly obtained through the frequency response data.
(2) A small-amplitude step signal is used to excite the PCA. Figure 6.12 (indicated by the solid blue line) shows the transient response of the PANS which is captured during the first 6 ms. With the identified resonant frequency ω_n and the gain k_p in the first step, the transient response data is used to adjust time constant τ and damping coefficient ς.

Based on the above two-step method, the identified parameters are compiled in Table 6.2. Therefore, the linear dynamic plant $G(s)$ of the proposed model is expressed as

$$G(s) = \frac{4.767 \times 10^6}{s^3 + 7582s^2 + 4.1 \times 10^8 s + 1.677 \times 10^{12}}. \tag{6.38}$$

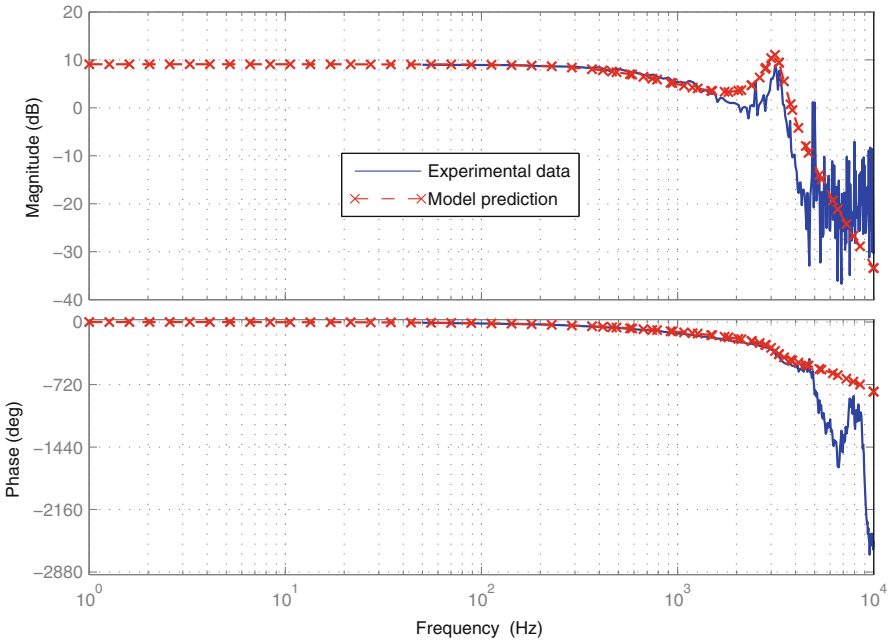


Fig. 6.11 Comparison of frequency response of the PANS (*blue solid*-experimental data; *red dash*-model simulation data)

6.5.3 *Identification of the Hysteresis Nonlinearity*

In order to identify the hysteresis nonlinearity in (6.34), the first step is to find a hysteresis model, for instance, the differential-equation-based or operator-based model (as detailed in Sect. 6.4), to describe the behaviors of the hysteresis. In this chapter, the MPI model [30] is selected as an illustration for hysteresis description. Certainly, other hysteresis models can also be selected. Without loss of generality, in what follows we use the MPI hysteresis model to validate the proposed model.

Before introducing the MPI model, the one-side play operator $F_{\text{or}}[v_{\text{in}}](t)$ with a threshold $r \geq 0$ for any piecewise monotone input function $v_{\text{in}}(t) \in C_m[0, t_E]$ is defined as

$$\begin{aligned} u(0) &= F_{\text{or}}[v_{\text{in}}](0) = f_{\text{or}}(v_{\text{in}}(0), 0) \\ u(t) &= F_{\text{or}}[v_{\text{in}}](t) = f_{\text{or}}(v_{\text{in}}(t), u(t_i)) \end{aligned} \tag{6.39}$$

for $t_i < t \leq t_{i+1}, 0 \leq i \leq N-1$ with

$$f_{\text{or}}(v_{\text{in}}, u) = \max(v_{\text{in}} - r, \min(v_{\text{in}}, u)) \tag{6.40}$$

for $0 = t_0 < t_1 < \ldots < t_N = t_E$.

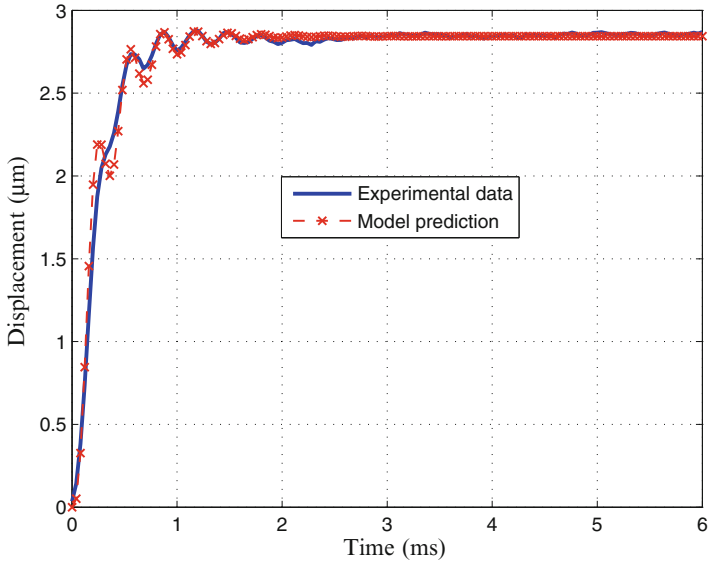


Fig. 6.12 Comparison of 1-V step response of the PANS in the first 6 ms (*blue solid*-experimental output response; *red dash*-model output response)

Table 6.2 Identified parameters of the general model for the PANS

Parameters	Value	Units
m	0.0515	kg
b_s	114	Ns/m
k_s	16.08	N/μm
R_0	40.5	Ω
C_A	4.6	μF
T_{em}	4.575	N/V
k_{amp}	10	–
a_0	7582	–
a_1	4.1×10^8	–
a_2	1.677×10^{12}	–
K	4.767×10^6	–

Then, with a finite number of the one-side play operators $F_{\text{or}}[v_{\text{in}}](t)$, the MPI model can be expressed in the form of Gu et al. [30]

$$w(t) = g(v_{\text{in}})(t) + \sum_{i=1}^{n} b_i F_{\text{or}_i}[v_{\text{in}}](t), \tag{6.41}$$

where $g(v_{\text{in}})(t) = p_1 v_{\text{in}}^3(t) + p_2 v_{\text{in}}(t)$ is a function of the input $v_{\text{in}}(t)$ with the constant parameters p_1 and p_2, n is the number of the adopted play operators for modeling, and $b_i = p(r_i)(r_i - r_{i-1})$ is the weighted coefficient for the threshold r_i, satisfying

$$r_i = \frac{i-1}{n}||v_{\text{in}}(t)||_\infty,\ i = 1, 2, \ldots, n. \tag{6.42}$$

For detailed descriptions of the MPI model, the reader may refer to [30].

With the MPI model for hysteresis description, it is ready to identify the parameters in (6.41) on the basis of the previous identified parameters of the linear dynamic plant (6.38). In fact, due to the highly nonlinear, high dimensional, and multiple constraint characteristics, identification of the hysteresis models is also a challenging task and many identification algorithms, such as the least square method, the genetic algorithms, and particle swarm optimization (PSO), have been reported to solve this problem [8, 46, 90]. PSO [90] has been shown to be superior to its competitors for identification of the P–I model. In this work, the effective informed adaptive PSO [47, 90] is borrowed to identify the MPI model.

From the MPI model described in (6.41), when ten one-side play operators are utilized, i.e., $n = 10$, it can be seen that there are 12 parameters to be identified, that is:

$$\mathbf{X} = [p_1, p_2, b_1, b_2, \ldots, b_{10}]. \tag{6.43}$$

As the parameters are difficult to find by trials, a minimize optimization is used to solve the problem. The key point of the optimization is the selection of the fitness function (or objective function), which has an important influence on the identified results. In this work, the fitness function is chosen as:

$$F(\mathbf{X}) = \frac{1}{N}\sum_{i=1}^{N}(x_i^a - x_i)^2 \tag{6.44}$$

satisfying

$$\begin{aligned} &\dddot{x}(t) + a_0\ddot{x}(t) + a_1\dot{x}(t) + a_2x(t) = Kw(t) \\ &w(t) = p_1v_{\text{in}}^3(t) + p_2v_{\text{in}}(t) + \textstyle\sum_{i=1}^{n} b_iF_{\text{or}_i}[v_{\text{in}}](t), \end{aligned} \tag{6.45}$$

where N denotes the total number of the experimental data, x_i and x_i^a are the experimental data and simulation data at the ith sampling time, respectively. When the simulation data x_i^a is very close to the experimental data x_i, the identified result $\mathbf{X}$ is thought to be a good result. To sum up, the flowchart of the PSO algorithm is given in Fig. 6.13. Table 6.3 lists the identified parameters of the MPI model (6.41). The detailed discussions on the PSO algorithms may refer to [47, 90].

Remark 4. It should be noted that the larger n is selected, it is more precision to describe the hysteresis loops in theory. In real applications, it is verified that further increase of the numbers of the play operators leads to no significant improvement in the modeling accuracy. As in [30], there is a comparison of identification errors with different values of n, where it can be seen that the identification errors are at the same level with the selected values of n. The reader may refer to [30] for a detailed discussion. Therefore, in this development, we select $n = 10$ for a case study.

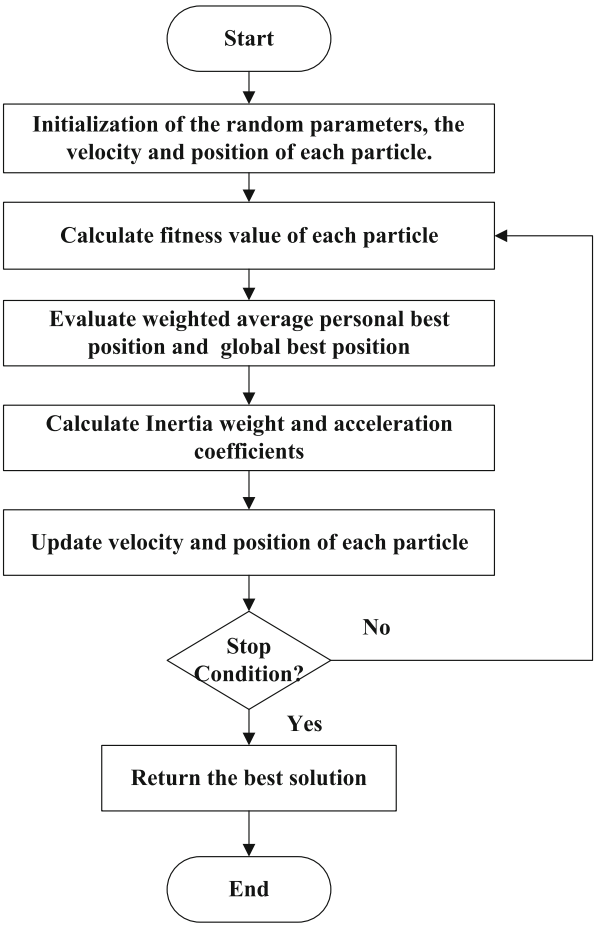


Fig. 6.13 The flowchart of the PSO identification algorithm

With the identified parameters of the linear dynamic plant and the nonlinear MPI hysteresis model, we are now ready to validate the proposed general model. Figure 6.14 shows the comparison of experimental data of the PANS and associated prediction results of the identified general model. In Fig. 6.14, the waveforms of the input voltage, displacement, prediction error, and hysteresis loops are shown.

From Fig. 6.14b, it can be obviously observed that the prediction results of the developed general model agree well with the experimental measurements. Figure 6.14c shows that the maximal prediction error is less than 1 % as a percentage of the full displacement range. In addition, the comparison of the hysteresis loops is given in Fig. 6.14d, where the asymmetric hysteresis nonlinearity of the PANS with the major and minor loops are well described by the developed model.

Table 6.3 Identified parameters of the MPI model

Number	r_i	b_i	p_i
1	0	0.0136	0.1189
2	0.1	0.0164	0.0422
3	0.2	0.0362	
4	0.3	0.0025	
5	0.4	0.0178	
6	0.5	0.0635	
7	0.6	0.0514	
8	0.7	0.0082	
9	0.8	0.1189	
10	0.9	0.0593	

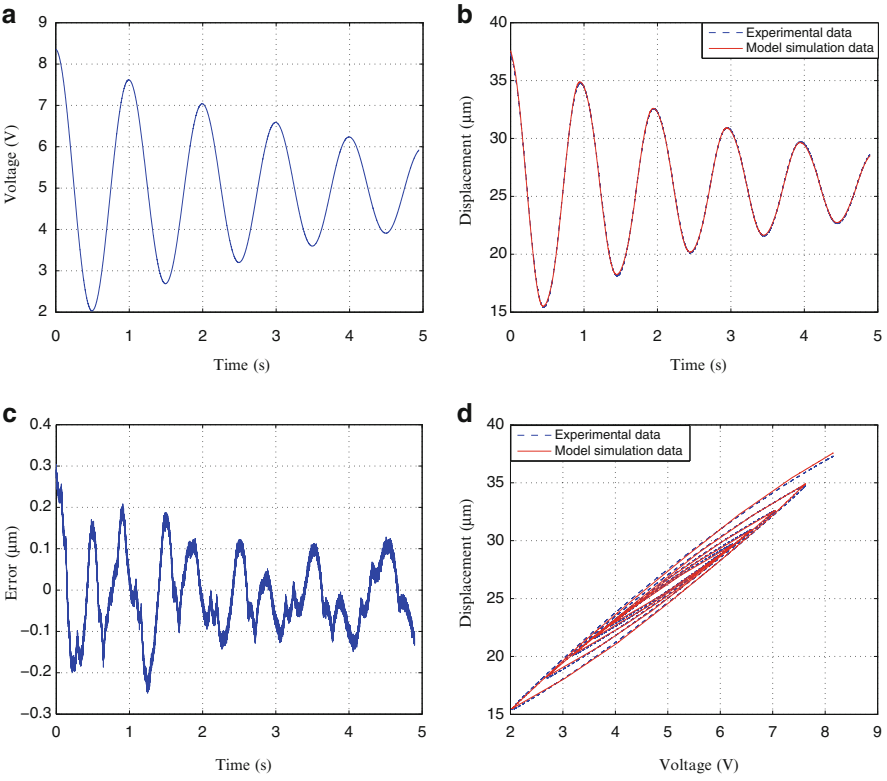


Fig. 6.14 Experimental validation of the proposed general model. (**a**) Input voltage. (**b**) Output displacement. (**c**) Predicted error. (**d**) Hysteresis loops

Therefore, from Figs. 6.11, 6.12 and 6.14, the proposed model with the identified parameters well characterize both the dynamics and hysteresis behaviors of the PANS, which clearly validates our developed general model.

6.6 Conclusion

PANSs are becoming promising in many precision manufacturing equipments. However, they suffer from the inherent hysteresis effect of the PCAs. The hysteresis effect exhibits complex nonlinear characteristics, which usually introduces undesirable inaccuracies or oscillations and even instability. Interest in studying PANSs with actuator hysteresis is also motivated by the fact that the hysteresis nonlinearity is coupled with the system dynamics for which traditional control methods are insufficient and thus require the comprehensive dynamic model of the PANS.

In this chapter, we first propose a general model to completely characterize nonlinear electric behavior, voltage-charge hysteresis, piezoelectric effect, and frequency response of the PANS in both electrical and mechanical aspects. After a brief survey of the hysteresis models, a MPI hysteresis model is utilized as an illustration to describe the hysteresis nonlinearity. To identify the parameters of the general model, the adopted approach is to express the general model into a third-order linear plant preceded by a MPI model. Then, the parameters of the linear plant and the hysteresis term are identified separately, which is confirmed by experimental results as well.

Acknowledgements This work was supported by the National Natural Science Foundation of China under Grant No. 51405293.

References

1. H. Adriaens, W.L. de Koning, R. Banning, Modeling piezoelectric actuators. IEEE/ASME Trans. Mechatron. **5**(4), 331–341 (2000)
2. H.T. Banks, A.J. Kurdila, G. Webb, Identification of hysteretic control influence operators representing smart actuators Part I: formulation. Math. Probl. Eng. **3**(4), 287–328 (1997a)
3. H.T. Banks, A.J. Kurdila, G. Webb, Identification of hysteretic control influence operators representing smart actuators Part II: convergent approximations. Math. Probl. Eng. **8**(6), 536–550 (1997b)
4. S. Bashash, N. Jalili, A polynomial-based linear mapping strategy for feedforward compensation of hysteresis in piezoelectric actuators. J. Dyn. Syst. Meas. Control **130**(3), 1–10 (2008)
5. S. Bashash, N. Jalili, Robust adaptive control of coupled parallel piezo-flexural nanopositioning stages. IEEE/ASME Trans. Mechatron. **14**(1), 11–20 (2009)
6. G. Bertotti, I.D. Mayergoyz, *The Science of Hysteresis: Mathematical Modeling and Applications*. (Academic, Oxford, 2006)
7. M. Brokate, J. Sprekels, *Hysteresis and Phase Transitions* (Springer, New York, 1996)
8. C. Chan, G. Liu, Hysteresis identification and compensation using a genetic algorithm with adaptive search space. Mechatronics **17**(7), 391–402 (2007)
9. X. Chen, T. Hisayam, Adaptive sliding-mode position control for piezo-actuated stage. IEEE Trans. Ind. Electron. **55**(11), 3927–3934 (2008)
10. S. Chonan, Z. Jiang, T. Yamamoto, Nonlinear hysteresis compensation of piezoelectric ceramic actuators. J. Intell. Mater. Syst. Struct. **7**(2),150–156 (1996)
11. B.D. Coleman, M.L. Hodgdon, On a class of constitutive relations for ferromagnetic hysteresis. Arch. Ration. Mech. Anal. **99**(4), 375–396 (1987)

12. J.M. Cruz-Hernandez, V. Hayward, Phase control approach to hysteresis reduction. IEEE Trans. Control Syst. Technol. **9**(1), 17–26 (2001)
13. S. Devasia, E. Eleftheriou, S.O.R. Moheimani, A survey of control issues in nanopositioning. IEEE Trans. Control Syst. Technol. **15**(5), 802–823 (2007)
14. R. Dong, Y. Tan, A modified Prandtl-Ishlinskii modeling method for hysteresis. Physica B **404**(8–11), 1336–1342 (2009)
15. A. Esbrook, X. Tan, H.K. Khalil, Control of systems with hysteresis via servocompensation and its application to nanopositioning. IEEE Trans. Control Syst. Technol. **21**(3), 725–738 (2013)
16. A.J. Fleming, A review of nanometer resolution position sensors: operation and performance. Sens. Actuators A Phys. **199**, 106–126 (2013)
17. A.J. Fleming, K.K. Leang, Charge drives for scanning probe microscope positioning stages. Ultramicroscopy **108**(12), 1551–1557 (2008)
18. W.S. Galinaitis, Two methods for modeling scalar hysteresis and their use in controlling actuators with hysteresis. Ph.D. thesis, Virginia Polytechnic Institute and State University (1999)
19. Y.S. Gao, D.W. Zhang, C.W. Yu, Dynamic modeling of a novel workpiece table for active surface grinding control. Int. J. Mach. Tools Manuf. **41**(4), 609–624 (2001)
20. H. Georgiou, R.B. Mrad, Electromechanical modeling of piezoceramic actuators for dynamic loading applications. J. Dyn. Syst. Meas. Contr. **128**(3), 558–567 (2006)
21. F. Giri, Y. Rochdi, F.Z. Chaoui, A. Brouri, Identification of Hammerstein systems in presence of hysteresis-backlash and hysteresis-relay nonlinearities. Automatica **44**(3), 767–775 (2008)
22. M. Goldfarb, N. Celanovic, Modeling piezoelectric stack actuators for control of micromanipulation. IEEE Control. Syst. Mag. **17**(3), 69–79 (1997)
23. G.Y. Gu, L.M. Zhu, High-speed tracking control of piezoelectric actuators using an ellipse-based hysteresis model. Rev. Sci. Instrum. **81**(8), 085104 (2010)
24. G.Y. Gu, L.M. Zhu, Modeling of rate-dependent hysteresis in piezoelectric actuators using a family of ellipses. Sens. Actuators A Phys. **165**(2), 202–209 (2011)
25. G.Y. Gu, L.M. Zhu, Motion control of piezoceramic actuators with creep, hysteresis and vibration compensation. Sens. Actuators A Phys. **197**, 76–87 (2013)
26. G.Y. Gu, M.J. Yang, L.M. Zhu, Real-time inverse hysteresis compensation of piezoelectric actuators with a modified Prandtl-Ishlinskii model. Rev. Sci. Instrum. **83**(6), 065106 (2012)
27. G.Y. Gu, L.M. Zhu, C.Y. Su, H. Ding, Motion control of piezoelectric positioning stages: modeling, controller design and experimental evaluation. IEEE/ASME Trans. Mechatron. **18**(5), 1459–1471 (2013)
28. G.Y. Gu, L.M. Zhu, C.Y. Su, High-precision control of piezoelectric nanopositioning stages using hysteresis compensator and disturbance observer. Smart Mater. Struct. **23**(10), 105007 (2014a)
29. G.Y. Gu, L.M. Zhu, C.Y. Su, Integral resonant damping for high-bandwidth control of piezoceramic stack actuators with asymmetric hysteresis nonlinearity. Mechatronics **24**(4), 367–375 (2014b)
30. G.Y. Gu, L.M. Zhu, C.Y. Su, Modeling and compensation of asymmetric hysteresis nonlinearity for piezoceramic actuators with a modified Prandtl-Ishlinskii model. IEEE Trans. Ind. Electron. **61**(3), 1583–1595 (2014c)
31. G.Y. Gu, L.M. Zhu, C.Y. Su, H. Ding, S. Fatikow, Modeling and control of piezo-actuated nanopositioning stages: a survey. IEEE Trans. Autom. Sci. Eng. (2014d). doi:10.1109/TASE.2014.2352364
32. G.Y. Gu, L.M. Zhu, C.Y. Su, H. Ding, S. Fatikow, Proxy-based sliding mode tracking control of piezoelectric-actuated nanopositioning stages. IEEE/ASME Trans. Mechatron. **20**(4), 1956–1965 (2015)
33. H. Habibullah, H.R. Pota, I.R. Petersen, M.S. Rana, Tracking of triangular reference signals using LQG controllers for lateral positioning of an AFM scanner stage. IEEE/ASME Trans. Mechatron. **19**(4), 1105–1114 (2014)

34. H. Hu, H. Georgiou, R.B. Mrad, Enhancement of tracking ability in piezoceramic actuators subject to dynamic excitation conditions. IEEE/ASME Trans. Mechatron. **10**(2), 230–239 (2005)
35. R.V. Iyer, X. Tan, Control of hysteretic systems through inverse compensation. IEEE Control. Syst. Mag. **29**(1), 83–99 (2009)
36. M.A. Janaideh, S. Rakheja, C.Y. Su, Experimental characterization and modeling of rate-dependent hysteresis of a piezoceramic actuator. Mechatronics **19**(5), 656–670 (2009)
37. M.A. Janaideh, S. Rakheja, C.Y. Su, An analytical generalized Prandtl-Ishlinskii model inversion for hysteresis compensation in micropositioning control. IEEE/ASME Trans. Mechatron. **16**(4), 734–744 (2011)
38. H. Janocha, K. Kuhnen, Real-time compensation of hysteresis and creep in piezoelectric actuators. Sens. Actuators A Phys. **79**(2), 83–89 (2000)
39. H. Jiang, H. Ji, J. Qiu, Y. Chen, A modified Prandtl-Ishlinskii model for modeling asymmetric hysteresis of piezoelectric actuators. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **57**(5), 1200–1210 (2010)
40. D.C. Jiles, D.L. Atherton, Theory of ferromagnetic hysteresis. J. Magn. Magn. Mater. **61**(1–2), 48–60 (1986)
41. T.L. Jordan, Z. Ounaies, Piezoelectric ceramics characterization. Technical report, ICASE, NASA Langley Research Center, Hampton (2001)
42. M. Krasnosel'skii, P. Pokrovskii, *Systems with Hysteresis* (Springer, Berlin, 1989)
43. P. Krejci, *Hysteresis Convexity and Dissipation in Hyperbolic Equations* (Gakkotosho, Tokyo, 1996)
44. P. Krejci, K. Kuhnen, Inverse control of systems with hysteresis and creep. IEE Proc. Control Theory Appl. **148**(3), 185–192 (2001)
45. K. Kuhnen, Modeling, identification and compensation of complex hysteretic nonlinearities: a modified Prandtl-Ishlinskii approach. Eur. J. Control **9**(4), 407–418 (2003)
46. K. Kuhnen, P. Krejci, Compensation of complex hysteresis and creep effects in piezoelectrically actuated systems-a new Preisach modeling approach. IEEE Trans. Autom. Control **54**(3), 537–550 (2009)
47. Q. Li, W. Chen, Y. Wang, S. Liu, J. Jia, Parameter identification for PEM fuel-cell mechanism model based on effective informed adaptive particle swarm optimization. IEEE Trans. Ind. Electron. **58**(6), 2410–2419 (2011)
48. C.X. Li, G.Y. Gu, M.J. Yang, L.M. Zhu, Design, analysis and testing of a parallel-kinematic high-bandwidth XY nanopositioning stage. Rev. Sci. Instrum. **84**(12), 125111 (2013a)
49. P. Li, F. Yan, C. Ge, X. Wang, L. Xu, J. Guo, P. Li, A simple fuzzy system for modelling of both rate-independent and rate-dependent hysteresis in piezoelectric actuators. Mech. Syst. Signal Process. **36**(1), 182–192 (2013b)
50. W. Li, X. Chen, Compensation of hysteresis in piezoelectric actuators without dynamics modeling. Sens. Actuators A Phys. **199**, 89–97 (2013)
51. Y. Li, Q. Xu, Adaptive sliding mode control with perturbation estimation and PID sliding surface for motion tracking of a piezo-driven micromanipulator. IEEE Trans. Control Syst. Technol. **18**(4), 798–810 (2010)
52. H.C. Liaw, B. Shirinzadeh, J. Smith, Enhanced sliding mode motion tracking control of piezoelectric actuators. Sens. Actuators A Phys. **138**(1), 194–202 (2007)
53. C.J. Lin, S.Y. Chen, Evolutionary algorithm based feedforward control for contouring of a biaxial piezo-actuated stage. Mechatronics **19**(6), 829–839 (2009)
54. C.J. Lin, P.T. Lin, Tracking control of a biaxial piezo-actuated positioning stage using generalized Duhem model. Comput. Math. Appl. **64**(5), 766–787 (2012)
55. J.W. Machi, P. Nistri, P. Zecca, Mathematical models for hysteresis. SIAM Rev. **35**(1), 94–123 (1993)
56. I. Mayergoyz, Preisach models of hysteresis. IEEE Trans. Magn. **24**(6), 2925–2927 (1988)
57. I. Mayergoyz, *Mathematical Models of Hysteresis and Their Applications* (Elsevier Academic Press, Amsterdam, 2003)

58. A. Meitzler, H. Tiersten, A. Warner, D. Berlincourt, G. Couqin, F. Welsh III, IEEE standard on piezoelectricity (1988)
59. J. Minase, T.F. Lu, B. Cazzolato, S. Grainger, A review, supported by experimental results, of voltage, charge and capacitor insertion method for driving piezoelectric actuators. Precis. Eng. **34**(4), 692–700 (2010)
60. M. Mohammed Ismail, F. Ikhouane, J. Rodellar, The hysteresis Bouc-Wen model, a survey. Arch. Comput. Methods Eng. **16**(2), 161–188 (2009)
61. R.B. Mrad, H. Hu, A model for voltage-to-displacement dynamics in piezoceramic actuators subject to dynamic-voltage excitations. IEEE/ASME Trans. Mechatron. **7**(4), 479–489 (2002)
62. C. Niezrecki, D. Brei, S. Balakrishnan, A. Moskalik, Piezoelectric actuation: state of the art. Shock Vib. Dig. **33**(4), 269–280 (2001)
63. J. Oh, D.S. Bernstein, Semilinear Duhem model for rate-independent and rate-dependent hysteresis. IEEE Trans. Autom. Control **50**(5), 631–645 (2005)
64. A. Padthe, B. Drincic, J. Oh, D. Rizos, S. Fassois, D. Bernstein, Duhem modeling of friction-induced hysteresis. IEEE Control Syst. **28**(5), 90–107 (2008)
65. G. Park, M.T. Bement, D.A. Hartman, R.E. Smith, C.R. Farrar, The use of active materials for machining processes: a review. Int. J. Mach. Tools Manuf. **47**(15), 2189–2206 (2007)
66. Y. Qin, B. Shirinzadeh, Y. Tian, D. Zhang, Design issues in a decoupled XY stage: static and dynamics modeling, hysteresis compensation, and tracking control. Sens. Actuators A Phys. **194**, 95–105 (2013)
67. M. Quant, H. Elizalde, A. Flores, R. Ramirez, P. Orta, G. Song, A comprehensive model for piezoceramic actuators: modelling, validation and application. Smart Mater. Struct. **18**(12), 125011 (2009)
68. M. Rakotondrabe, Bouc-Wen modeling and inverse multiplicative structure to compensate hysteresis nonlinearity in piezoelectric actuators. IEEE Trans. Autom. Sci. Eng. **8**(2), 428–431 (2011)
69. B. Ren, P. San, S. Ge, T.H. Lee, Adaptive dynamic surface control for a class of strict-feedback nonlinear systems with unknown backlash-like hysteresis, in *American Control Conference*, pp. 4482–4487 (2009)
70. L. Riccardi, D. Naso, B. Turchiano, H. Janocha, Design of linear feedback controllers for dynamic systems with hysteresis. IEEE Trans. Control Syst. Technol. **22**(4), 1268–1280 (2014)
71. S. Rosenbaum, M. Ruderman, T. Strohla, T. Bertram, Use of Jiles-Atherton and Preisach hysteresis models for inverse feed-forward control. IEEE Trans. Magn. **46**(12), 3984–3989 (2010)
72. C. Ru, L. Sun, Hysteresis and creep compensation for piezoelectric actuator in open-loop operation. Sens. Actuators A Phys. **122**(1), 124–130 (2005)
73. M. Ruderman, Direct recursive identification of the preisach hysteresis density function. J. Magn. Magn. Mater. **348**(12), 22–26 (2013)
74. M. Ruderman, T. Bertram, Discrete dynamic Preisach model for robust inverse control of hysteresis systems, in *49th IEEE Conference on Decision and Control*, pp. 3463–3468 (2010)
75. S.M. Salapaka, M.V. Salapaka, Scanning probe microscopy. IEEE Control Syst. Mag. **28**(2), 65–83 (2008)
76. Y. Shan, K.K. Leang, Accounting for hysteresis in repetitive control design: nanopositioning example. Automatica **48**(4), 1751–1758 (2012)
77. J. Shen, W.Y. Jywe, H.K. Chiang, Y.L. Shu, Precision tracking control of a piezoelectric-actuated system. Precis. Eng. **32**(2), 71–78 (2008)
78. R. Smith, Z. Ounaie, A domain wall model for hysteresis in piezoelectric materials. J. Intell. Mater. Syst. Struct. **11**(1), 62–79 (2000)
79. D. Song, C.J. Li, Modeling of piezo actuator's nonlinear and frequency dependent dynamics. Mechatronics **9**(4), 391–410 (1999)
80. G. Song, J.Q. Zhao, X.Q. Zhou, J.A.D. Abreu-Garcia, Tracking control of a piezoceramic actuator with hysteresis compensation using inverse Preisach model. IEEE/ASME Trans. Mechatron. **10**(2), 198–209 (2005)

81. C.Y. Su, Y. Stepanenko, J. Svoboda, T.P. Leung, Robust adaptive control of a class of nonlinear systems with unknown backlash-like hysteresis. IEEE Trans. Autom. Control **45**(12), 2427–2432 (2000)
82. U.X. Tan, T.L. Win, W.T. Ang, Modeling piezoelectric actuator hysteresis with singularity free Prandtl-Ishlinskii model, in *IEEE International Conference on Robotics and Biomimetics - ROBIO2006*, pp. 251–256 (2006)
83. G. Tao, F.L. Lewis, *Adaptive Control of Nonsmooth Dynamic Systems* (Springer, New York, 2001)
84. A. Visintin, *Differential Models of Hysteresis* (Springer, Berlin, 1994)
85. C. Visone, Hysteresis modelling and compensation for smart sensors and actuators. J. Phys. Conf. Ser. **138**(1), 012028 (2008)
86. S.R. Viswamurthy, R. Ganguli, Modeling and compensation of piezoceramic actuator hysteresis for helicopter vibration control. Sens. Actuators A Phys. **135**(2), 801–810 (2007)
87. G.V. Webb, D.C. Lagoudas, A.J. Kurdila, Hysteresis modeling of SMA actuators for control applications. J. Intell. Mater. Syst. Struct. **9**(6), 432–448 (1998)
88. Q. Xu, Y. Li, Micro-/nanopositioning using model predictive output integral discrete sliding mode control. IEEE Trans. Ind. Electron. **59**(2), 1161–1170 (2012)
89. Q. Xu, P.K. Wong, Hysteresis modeling and compensation of a piezostage using least squares support vector machines. Mechatronics **21**(7), 1239–1251 (2011)
90. M.J. Yang, G.Y. Gu, L.M. Zhu, Parameter identification of the generalized Prandtl-Ishlinskii model for piezoelectric actuators using modified particle swarm optimization. Sens. Actuators A Phys. **189**, 254–265 (2013)
91. J. Yi, S. Chang, Y. Shen, Disturbance-observer-based hysteresis compensation for piezoelectric actuators. IEEE/ASME Trans. Mechatron. **14**(4), 456–464 (2009)
92. Y.H. Yu, Z.C. Xiao, N.G. Naganathan, R.V. Dukkipat, Dynamic Preisach modeling of hysteresis for the piezoceramic actuator system. Mech. Mach. Theory **37**(1), 75–89 (2002)
93. X. Zhao, Y. Tan, Modeling hysteresis and its inverse model using neural networks based on expanded input space method. IEEE Trans. Control Syst. Technol. **16**(3), 484–490 (2008)
94. J.H. Zhong, B. Yao, Adaptive robust precision motion control of a piezoelectric positioning stage. IEEE Trans. Control Syst. Technol. **16**(5), 1039–1046 (2008)
95. J. Zhou, C. Wen, C. Zhang, Adaptive backstepping control of a class of uncertain nonlinear systems with unknown backlash-like hysteresis. Automatica **49**(10), 1751–1759 (2004)

Chapter 7
Tracking Control for Nanopositioning Systems

Kam K. Leang and Andrew J. Fleming

Abstract The performance of nanopositioning systems is greatly affected by their mechanical dynamics, and for piezo-actuated designs, induced structural vibration, hysteresis, and creep can drastically limit positioning precision. Therefore, tracking control, both feedback and feedforward control, plays an important role in achieving high-performance operation, especially at high operating frequencies. This chapter reviews popular feedback and feedforward control techniques for nanopositioning systems. First, the effects of vibration, hysteresis, and creep are described, where simple methods traditionally employed to avoid these effects are discussed. Second, various models for nanopositioning systems for control system design, simulation, and synthesis are presented. Finally, popular feedback and feedforward controllers to handle vibration, hysteresis, and creep are presented, along with experimental results.

7.1 Introduction

Nanopositioning systems rely on tracking controllers for optimum performance. Extensive studies have been done to explore control techniques to compensate for the effects of vibration, hysteresis, and creep in nanopositioning systems [1–4]. In general, these techniques can be categorized as either feedback control or feedforward control, each with its advantages and disadvantages. In some cases, integrated feedback and feedforward controllers are employed.

Feedback control schemes (see basic block diagram in Fig. 7.1a) such as proportional-integral-derivative (PID) [5], state-feedback [6], and H_∞ control [7, 8] have demonstrated substantial reduction of positioning errors due to creep and hysteresis. The advantages of feedback control include: (1) the ability to handle modeling errors and (2) robustness to parameter variation, such as those caused by

K.K. Leang (✉)
University of Utah, Salt Lake City, UT, USA
e-mail: kam.k.leang@utah.edu

A.J. Fleming
University of Newcastle, Callaghan, Australia
e-mail: andrew.fleming@newcastle.edu.au

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_7

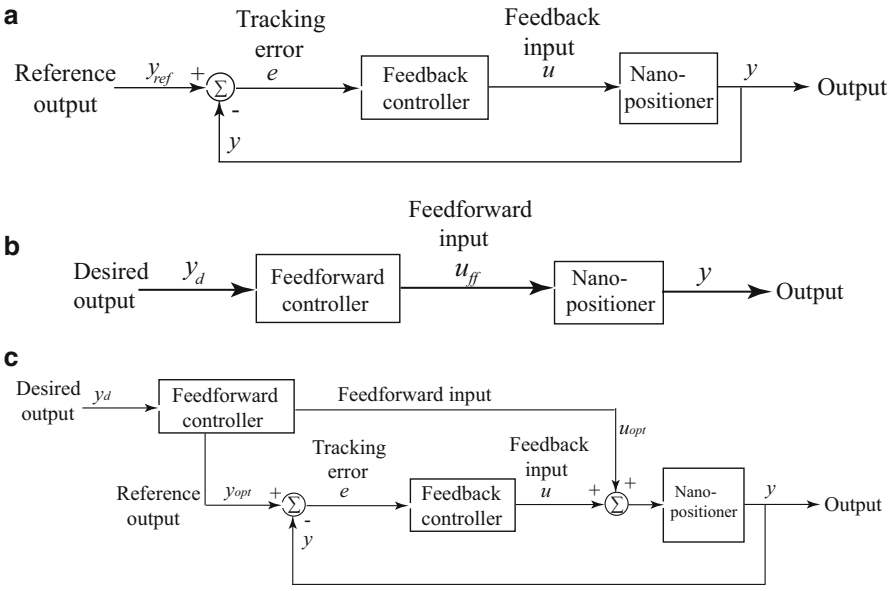


Fig. 7.1 Control systems for nanopositioning: (**a**) feedback controller, (**b**) feedforward controller, and (**c**) integrated feedback and feedforward controller

aging effects [9] and environmental changes (temperature changes [10]). Although feedback control reduces positioning errors, the low gain margin in piezo positioners tend to limit the achievable performance, i.e., high feedback gain tends to destabilize the system [5, 11]. In particular, low structural damping in piezo positioners results in high quality factor Q (i.e., sharp resonant peak) which gives rise to low gain margin. In practice, a compromise is sought between performance and instability; feedback gains (such as proportional, derivative, and integral terms) are adjusted to improve performance without instability (e.g., in references [12, 13]). Moreover, feedback-based approaches have had limited success in compensating for dynamics effects [5].

To overcome some of the performance limitations inherent in feedback control (e.g., limited dynamic compensation due to low gain margin), model-based feedforward control (see Fig. 7.1b) has been used to compensate for creep and hysteresis effects [14–19]. The feedforward input is found by carefully modeling the complicated creep and hysteresis behaviors and inverting the models to determine the control inputs. Many model-based approaches have been considered, for instance, Smith et al. [20] developed a hysteresis model based on energy methods and then used the model for the design of an inverse compensator. Croft et al. [21, 22] and Stilson et al. [23] compensated for all three effects using an inversion-based feedforward technique and they applied their technique to scanning probe microscopy (SPM) imaging. However, a major disadvantage of using model-based feedforward approaches to compensate for creep and hysteresis is that these

effects are difficult to model accurately and the approach is prone to errors [24]. Furthermore, model-based approaches can be computationally cumbersome to implement.

Finally, integrated feedback and feedforward control schemes (see Fig. 7.1c) have been developed [4, 25] to combine the advantages of both approaches. In particular, feedback control minimizes the nonlinear effects such as hysteresis and the operating speed is improved by the addition of feedforward input to compensate for the vibrational dynamics. This chapter presents popular feedback and feedforward controllers to handle vibration, hysteresis, and creep, along with experimental results.

7.2 Challenges in Tracking Control of Nanopositioning Systems

The performance of a nanopositioning system is affected by induced structural vibration (vibrational dynamics), hysteresis, and creep in the piezoactuator, cross-coupling behaviors in both the actuator and mechanical system, external disturbances, and drift due to temperature variations [2, 4, 26, 27]. Cross-coupling effects can be handled using multi-input-multi-output controllers and iterative methods [28]. External disturbances are often dealt with using passive or active vibration isolation systems. Below, the effects of vibrational dynamics, hysteresis, and creep are discussed, as well as their impact on positioning performance. Afterwards, tracking controllers to deal with these behaviors are presented.

7.2.1 Vibrational Dynamics

Vibration effects are caused by exciting the resonance modes of the positioning system [22, 29] and in general vibration limits the operating bandwidth. For example, the frequency response of a piezo-based positioner, such as the measured response for a tube-shaped piezoactuator shown in Fig. 7.2, typically reveals sharp resonant peaks. These peaks can easily be excited by certain command signals like triangle inputs applied to control the positioner. Figure 7.3 shows a measured displacement versus time response for a piezo-positioning system in which the three effects of vibration, hysteresis, and creep are observable. In particular, the effect of vibration leads to oscillations, which cause significant tracking error in the displacement versus time response. The slow drift in the response is caused by creep, and the curved appearance in the output response is due to hysteresis. Overall, such effects cause distortion in nanopositioning applications, including image distortion in atomic force microscope images (AFMs). In fact, vibrations cause ripples to appear in AFM images during high-speed scanning [30].

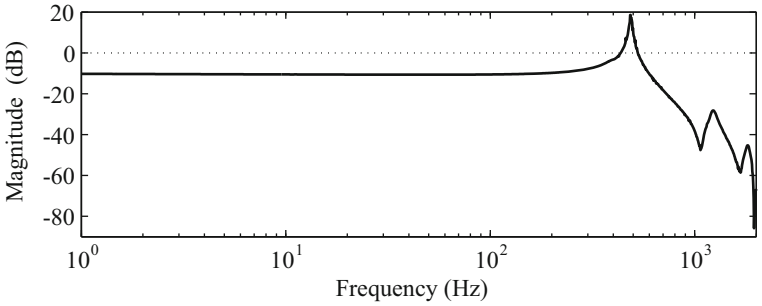


Fig. 7.2 Frequency response of a tube-shaped piezoactuator

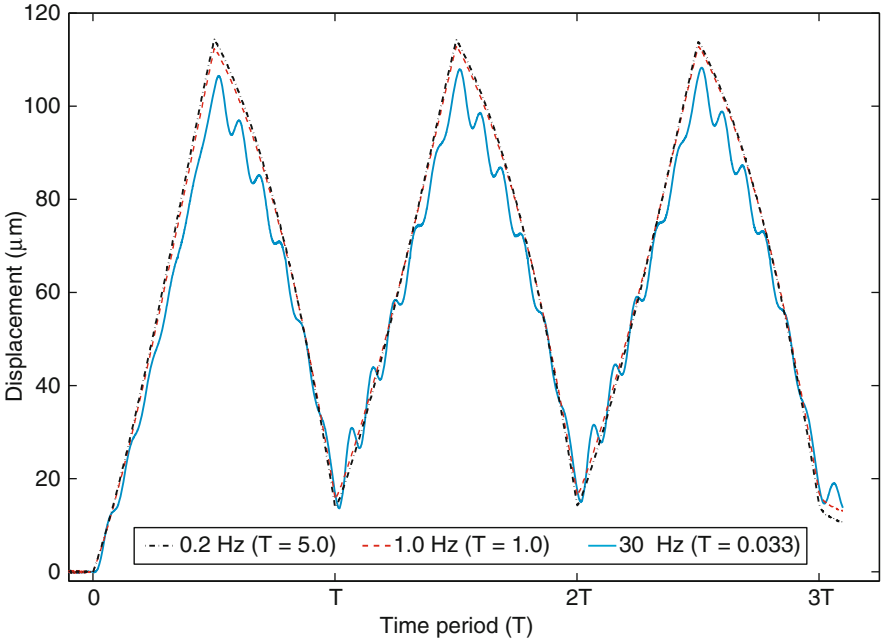


Fig. 7.3 Measured piezoactuator response (scanning at 30 Hz) showing the combined effects of vibration, hysteresis, and creep

Typically, scan rates (i.e., scan frequencies) are restricted to less than 1/100th to 1/10th of the first resonant frequency to minimize the effects of vibration, thus limiting the bandwidth of piezo-based systems because the achievable scan rate is lower for increased resolution in positioning. However, higher operating speed can be achieved by using *stiffer* piezoactuators with higher resonant frequencies [31–33], for example, Ando et al. [34] used a *stiff* piezo with a resonant frequency of 260 kHz in an AFM to image biological macromolecules in action. It is pointed

out that compact and stiff piezoactuators have shorter effective displacement ranges. Therefore, the use of stiffer piezos to increase bandwidth also leads to reduction of positioning range.

7.2.2 Hysteresis

Hysteresis, which is a nonlinear behavior between the applied electric field and the mechanical displacement of a piezoelectric actuator, is believed to be caused by irreversible losses that occur when similarly oriented electric dipoles interact upon application of an electric field [35]. The effect of hysteresis on the displacement of a piezoelectric actuator is more pronounced over large-range motion [5], as indicated in Fig. 7.3.

The term hysteresis comes from the Greek word "to be late" or "come behind" and it was first coined for application in 1881 by physicist Ewing when he was studying magnetization. Interestingly, the year 1881 was when the Curie Brothers were credited with the discovery of the piezoelectric effect. Hysteresis is often referred to as a lag in the response. An interesting writing on the history of hysteresis can be found in reference [36], which describes other systems which exhibit this behavior.

The mechanism responsible for hysteresis in piezoelectric transducers is better understood by considering the domain wall analogy for describing hysteresis in magnetic materials [35, 37]. For example, magnetic materials consist of tiny elementary magnetic dipoles. These particles align to an applied field. The analogy to this in piezoelectric materials the unit cell of the piezoelectric crystal material exhibits an electric dipole. The term *domains of polarization* refers to regions of similarly-oriented dipoles, that is, a relatively large region of connected unit cells having similarly-oriented net polarization. The imaginary boundary which separates these regions are referred to as *domain walls*. These boundaries grow or shrink depending on the nature of the applied field. For the simple case, an isolated elementary dipole subjected to an applied field will orient itself to the field instantaneously, and therefore displays no hysteresis. However, hysteresis is said to arise due to "internal forces," which cause the dipoles to exhibit a preference for their orientation, and hence the motion of the domain walls are retarded by such forces. These internal forces are attributed to material defects and internal friction between dipoles and the domain walls. Even though the domain wall analogy was conceived for magnetic materials, it can easily be extended to materials which consist of elementary dipole-like particles, such as piezoelectric materials. Additionally, hysteresis "remembers" the effect of the past, which further complicates the problem in terms of precision control.

Hysteresis in piezo-based nanopositioning systems causes AFM image distortion [5, 30], instability of the closed-loop [11], and loss in calibration [19, 38].

By operating the nanopositioner over a relatively small range, say less than 10 % of full range, hysteresis can be avoided. However, operating over such a narrow range limits the ability of the actuator for long-range motion with sub-nanometer resolution.

7.2.3 Creep

During slow and static positioning, creep and drift effects can cause over 30 % error in positioning [39]. Creep is a low-frequency behavior, where its time constant is typically a few minutes. Creep severely degrades the low-frequency and static positioning ability of piezoelectric actuators [40–42].

In mechanics, creep is a rate-dependent deformation of the material when subjected to a constant load or stress [43]. Similarly, creep in piezoelectric materials is a rate-dependent deformation due to a constant electric field. Creep manifests itself as the remnant polarization slowly increases after the onset of a constant field.

Figure 7.4 shows the effect of creep from an experimental piezoactuator. The actuator is commanded to a reference position, say 25 μm, but after a period of 15 min, the actuators' position creeps to a new position of 33.41 μm. As a result, the error due to creep is 24.44 % of the total displacement range. To avoid creep, the nanopositioner can be operated at a relatively fast rate (above 1 Hz scanning frequency) [30]; however, such effort prevents the use of piezo positioners in slow and static applications. For example, because of creep, it is difficult to precisely fabricate nanofeatures using AFMs when the process time-scale is on the order of minutes, e.g., see [40].

Methods to compensate for creep have been well studied in the past and some examples include the use of feedback control, e.g., [5, 8, 13, 44], and model-based feedforward control, e.g., [16, 19, 30, 42, 45, 46].

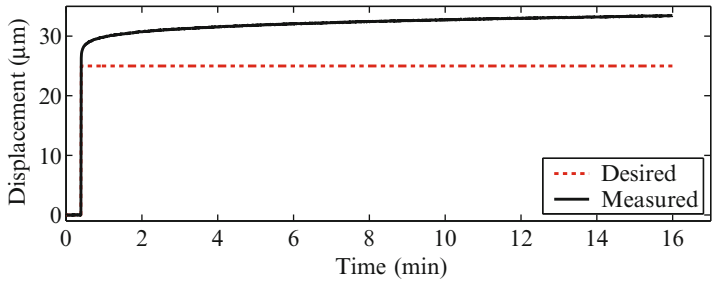


Fig. 7.4 The effects of creep in the output displacement measured over a period of 15 min

7.3 Modeling Vibration, Hysteresis, and Creep for Control

The displacement (stoke) response of a piezoelectric actuator employed in a conventional nanopositioning system under an applied voltage where the effects of temperature variations and self-heating are ignored consists of dynamics (vibration and creep) and nonlinearities (such as hysteresis). An overview of modeling of a piezo-based nanopositioner is considered. The main objective is to describe a modeling approach which enables the application of feedforward as well as feedback control for precision tracking control.

As illustrated in Fig. 7.3, the measured output response of a piezo-based nanopositioner consists of the combined effects of vibration, hysteresis, and creep. In Fig. 7.3, the oscillations are caused by vibrational dynamics; the slow upward-drift of the output over time is due to the creep effect; and finally, a noticeable curved distortion in the output trajectory is due to hysteresis.

To effectively model these behaviors, one must consider the operating conditions and when certain effects dominate. For instance, the amount of the dynamics on the output response depends on the operating frequency. The operating frequency is determined by the application in mind. As a result, the system's *dynamics*, be it low-frequency or high-frequency, are excited by the frequency of the input signal.

Take for example when the input frequency is close to a nanopositioning system's resonance frequencies. In this case, vibration becomes noticeably large. This is not surprising when one examines the frequency response of a typical piezoelectric actuator, where sharp resonances are common as shown in Fig. 7.2. Piezoactuators tend to be highly resonant structures due to their high stiffness and low structural damping. Therefore, input signals such as sawtooth signals can excite the piezoactuator's resonances, causing the output to oscillate or vibrate as previously shown in Fig. 7.3.

At slow operating speeds, creep is a major source of positioning error. Creep in piezoactuators is a low-frequency behavior, where the output drifts, especially when the operation is offset from the center of the piezoactuator's positioning range (see Fig. 7.3).

Finally, hysteresis is significant over large-range displacements [5]. The operation of piezoactuators in their linear range helps avoid hysteresis. In general, the linear range is within 5 to 10% of the maximal range of motion.

The hysteresis and dynamic effects are coupled [30]. For instance, when the movement of the piezoactuator is large and slow, the piezoactuator exhibits hysteresis and creep effects. As the input frequency increases, the piezoactuator's output response shows the addition of the vibrational dynamics. To model these behaviors, the cascade model depicted in Fig. 7.5a is used. The range-dependent hysteresis effect is treated as a rate-independent, input nonlinearity represented by $H[u(\cdot)]$. The vibrational dynamics and creep effects are captured by the linear dynamics model $G(s)$. The cascade model structure is used extensively to model piezoactuators and similar systems [30, 47]. Below, details on how to obtain models for these behaviors are discussed.

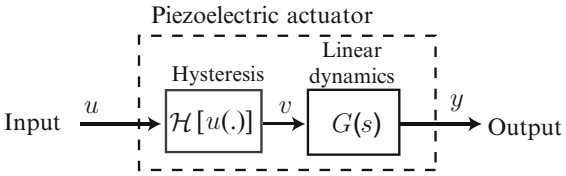


Fig. 7.5 A cascade model structure for hysteresis, vibrational dynamics, and creep effects in piezoactuators. Hysteresis, denoted by $H[u(\cdot)]$, is modeled as an rate-independent, input nonlinearity that is output-range dependent. The input-frequency-dependent vibrational dynamics and creep effects follow the hysteresis block, and they are captured by the linear dynamics model $G(s)$

7.3.1 Modeling Linear Dynamics: Vibration and Creep

The transfer function models for the vibrational dynamics and creep effects of a piezoactuator can be obtained in a number of ways. One straightforward approach is curve fitting the measured frequency and time responses over appropriate frequency ranges. Another is developing a model using white noise data or a chirp signal. The former approach is discussed below to model vibrational dynamics and creep.

To find the vibrational dynamics model, which relates the input u to the displacement in the x, y, or z axis, a system identification algorithm is applied to the measured frequency response. Consider a tube-shaped piezoactuator with transverse range of approximately 100 μm. First, the frequency response along the x-axis is measured over a displacement range of less than 10 % of the maximal range to avoid the hysteresis effect. To avoid creep, the response is measured over a wide frequency range, in this case, 1 Hz to 2 kHz. Frequency responses can be obtained using commercially available dynamic signal analyzers (such as Stanford Research Systems SRT785) or other system identification approaches.

Consider the example measured frequency response for a piezoactuator shown in Fig. 7.6, where the solid line is the measured response.

Using a system identification algorithm, such as the function "invfreqs" in Matlab, a transfer function model is fitted to the measured response. The dash line shown in Fig. 7.6 is the model given by

$$G_v(s) = \frac{7.2 \cdot 10^{13} s^2 + 2.3 \cdot 10^{16} s + 3.2 \cdot 10^{21}}{s^6 + 1.1 \cdot 10^4 s^5 + 9.5 \cdot 10^7 s^4 + 7.0 \cdot 10^{11} s^3 + 2.0 \cdot 10^{15} s^2 + 5.6 \cdot 10^{18} s + 1.0 \cdot 10^{22}}. \tag{7.1}$$

At very low speed, the creep effect is significant. This effect can be captured by the Kelvin-Voigt model, which consists of spring (k_i) and damper (c_i) elements [16, 48]. The lumped-parameter model shown in Fig. 7.7 is linear, and its transfer function is

$$G_c(s) = \frac{x(s)}{u(s)} = \frac{1}{k_0} + \sum_{i=1}^{n} \frac{1}{s c_i + k_i}, \tag{7.2}$$

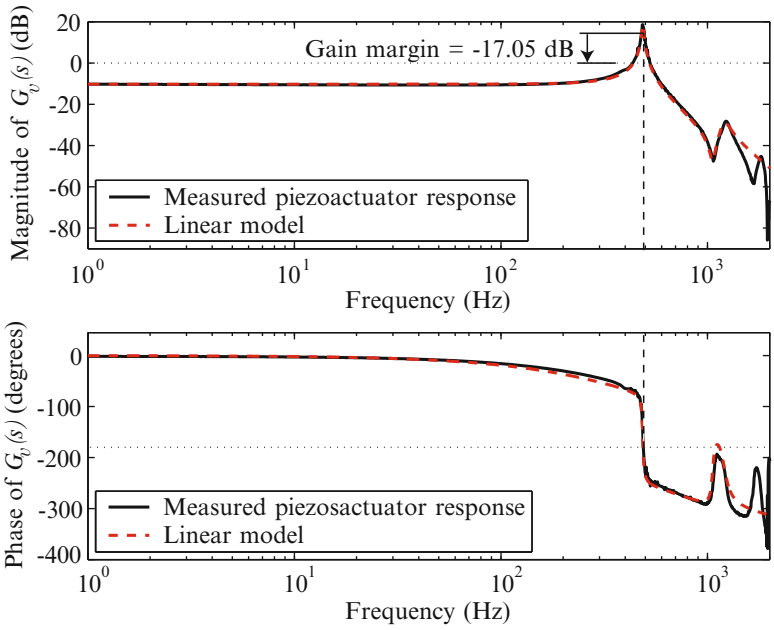


Fig. 7.6 The frequency response of a piezoactuator for modeling the vibrational dynamics. These plots show the measured frequency response (*solid line*), the magnitude and phase versus frequency, of the piezoactuator over ±2-μm displacement range. The *dash line* is the linear vibrational dynamics model

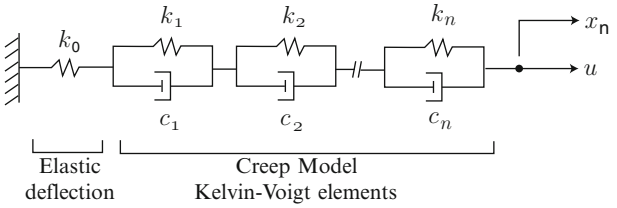


Fig. 7.7 The spring-damper model for the creep effect. The parameters of the model are found by curve fitting the measured step response of the piezoactuator

where $x(s)$ is the displacement of the piezoactuator and $u(s)$ is the applied input voltage. In (7.2), k_0 models the elastic behavior at dc, and the creep behavior is captured by selecting an appropriate model order corresponding to the number of spring-damper elements n. The parameters k_0, k_i, and c_i of (7.2) are often determined by curve fitting the step response of the piezoactuator measured over a long period of time.

The creep effect can also be modeled by a logarithmic model of the form

$$x(t) = x_0 \left[1 + \gamma \log_{10} \left(\frac{t}{t_0} \right) \right], \tag{7.3}$$

where $x(t)$ is the actuator's displacement, x_0 is the nominal fast displacement to the applied voltage, γ is the creep rate, and t_0 is the settling time of the high-frequency dynamics [19].

7.3.2 Modeling Hysteresis

Hysteresis is widely accepted as a nonlinear behavior characterized by a nonvanishing input-output loop as the frequency decreases to zero. The behavior sometimes corresponds to energy loss. Below are popular models for hysteresis for designing and synthesizing controllers.

7.3.2.1 Simple Polynomial Model

The simplest model is a polynomial fit of the output versus input data, with the form,

$$y = a_n u^n + a_{n-1} u^{n-1} + \cdots + a_1 u + a_0, \tag{7.4}$$

where n is the order of the polynomial, y is the output, u is the input, and the a_i's coefficients can be found using a least-squares algorithm. It is pointed out that this model ignores the branching (looping) behavior, that is, the ascending and descending branches are the same. Such a model provides reasonable accuracy for electrostrictive actuators as they tend to exhibit significantly less hysteresis compared to piezoelectric actuators [49]. Inversion of the polynomial hysteresis model for feedforward control is straightforward. A look-up table can be used in this case.

7.3.2.2 Maxwell Slip Model

A more realistic representation is the Maxwell slip model, a lumped-parameter approach [50]. In this model, the nonlinear input-output map and looping behavior are caused by a parallel combination of elasto-slide elements as shown in Fig. 7.8. An elasto-slide element consists of a massless linear spring and a massless block that is subjected to Coulomb friction as shown in Fig. 7.8. The constitutive behavior of the ith element is given by

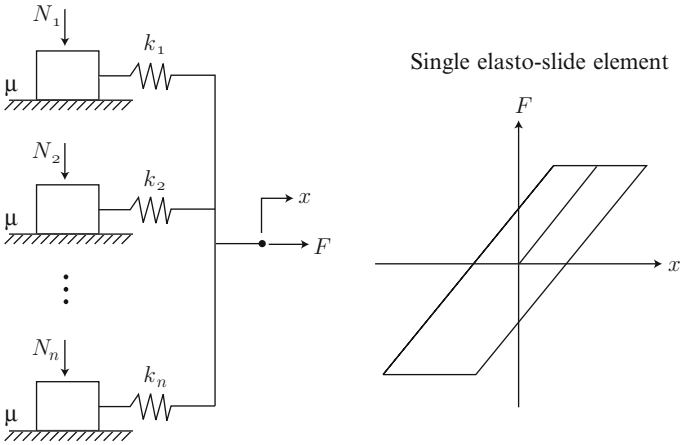


Fig. 7.8 Maxwell slip model for hysteresis

$$F_i = \begin{cases} k_i(x - x_{bi}) & \text{if } |k_i(x - x_{bi})| < f_i \\ f_i \text{sgn}(\dot{x}) & \text{and } x_{bi} = x - \frac{f_i}{k_i}\text{sgn}(\dot{x}) \text{ else,} \end{cases} \tag{7.5}$$

where x is the input displacement, F_i, k_i, f_i, and x_{bi} are the output force, spring stiffness, breakaway force, and block position, respectively, for the ith element. The resultant output force is computed as

$$F = \sum_{i=1}^{n} F_i. \tag{7.6}$$

Starting in a relaxed state, the measured input-output map can be used to determine the individual spring constants and break away forces (μ's). The Maxwell slip model has been applied to model and control a piezoelectric stack actuator in [50].

7.3.2.3 Duhem Model

A finite-dimensional, differential model typically used to model ferromagnetically soft materials [51] can be adapted to model the rate-independent hysteresis in the piezoactuator. Such a model is given by

$$\dot{v}(t) = \alpha|\dot{u}(t)|\big[\beta u(t) - v(t)\big] + \gamma\dot{u}(t), \tag{7.7}$$

where $v(t)$ is the output, $u(t)$ is the input, and α, β, γ are positive constants. The parameters of the model can obtained by a least-squares fit of the measured

input-output data. For example, the input $u(t) = A\sin(\omega t)$, with $\omega = 2\pi$ rad/s and $A = 2$ V, was applied to a piezoactuator and the output $v(t)$ was measured. Assuming zero initial conditions, Eq. (7.7) was rewritten in the following matrix form

$$\mathbf{v} = \mathbf{Q}\Theta, \tag{7.8}$$

where $\mathbf{v} = [v(t_1),\ v(t_2),\ v(t_3),\ \ldots]^T$ is the vector of measured outputs at specific time instances; $\Theta = [\alpha\beta,\ -\alpha,\ \gamma]^T$ is the vector of unknown parameters; and

$$\mathbf{Q} = \begin{bmatrix} q_1(t_1) & q_2(t_1) & q_3(t_1) \\ q_1(t_2) & q_2(t_2) & q_3(t_2) \\ q_1(t_3) & q_2(t_3) & q_3(t_3) \\ \vdots & \vdots & \vdots \end{bmatrix}, \tag{7.9}$$

where $q_3(t) = u(t)$ and

$$q_1(t) = \int |\dot{u}(t)|u(t)dt;\ \ q_2(t) = \int |\dot{u}(t)|v(t)dt. \tag{7.10}$$

The identified parameters were $\alpha = 0.3$, $\beta = 1.2$, and $\gamma = 1.0$. Figure 7.9a compares the normalized measured and model outputs for $u(t) = 2\sin(2\pi t)$ V.

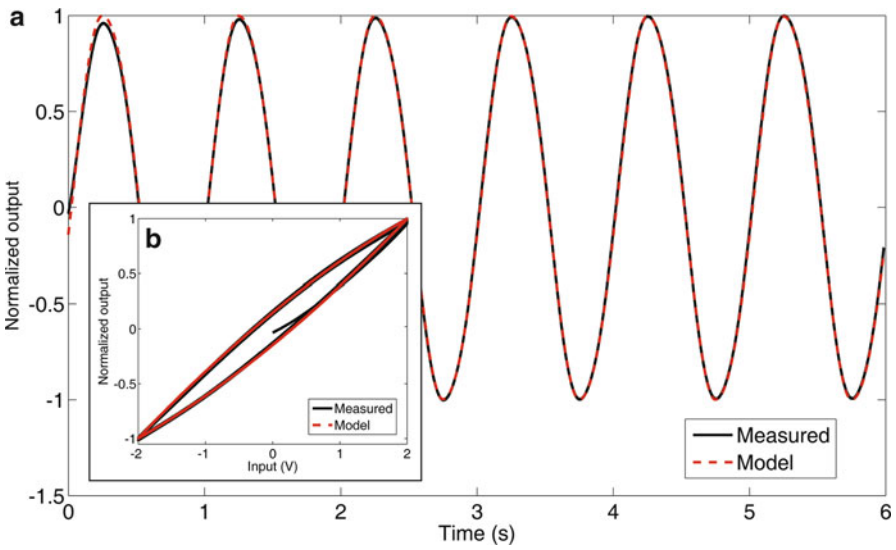


Fig. 7.9 Comparison of measured and Duhem hysteresis model output for a piezoactuator: (**a**) The normalized outputs versus time and (**b**) a hysteresis curve. Input was $u(t) = 2\sin(\omega t)$ V, with $\omega = 2\pi$ rad/s

The hysteresis curves are compared in Fig. 7.9b. The results show good agreement between the measured and model outputs, where the maximum steady-state error was less than 2 % of the total range.

7.3.2.4 Preisach Model

A relatively accurate hysteresis model is the Preisach hysteresis model, which was first developed in 1935 for magnetic materials [52]. This model has been studied extensively to characterize the rate-independent hysteresis in piezoelectric materials [53, 54], as well as many hysteretic systems, such as shape memory alloy devices [18]. This model can be inverted for feedforward control [30].

The Preisach model is a phenomenological description, whereby the output of a hysteretic system is the net effect of elementary relays, which represent the behavior of individual domains within the material. These domains or relays can assume a value of $+1$ or -1 depending on the current and future values of the input. The relay operator $R : R \to \{-1, +1\}$ is defined as [54]:

$$R_{\alpha,\beta}[u](t) = \begin{cases} +1 & u(t) > \alpha, \\ -1 & u(t) < \beta, \\ \text{unchanged} & \beta \leq u(t) \leq \alpha, \end{cases} \tag{7.11}$$

where $u(t)$ is the input. The pair (α, β) in Eq. (7.11), such that $\alpha \geq \beta$, represents the "up" and "down" switching values of the relay, respectively. Figure 7.10a shows an example of an elementary relay (also called Preisach hysteron).

The output $v(t)$, which is an infinite sum of hysterons [Fig. 7.10b], is written as

$$v(t) = H[u](t) = \iint_{\alpha \geq \beta} \mu(\alpha,\ \beta) R_{\alpha,\beta}[u](t) d\alpha d\beta, \tag{7.12}$$

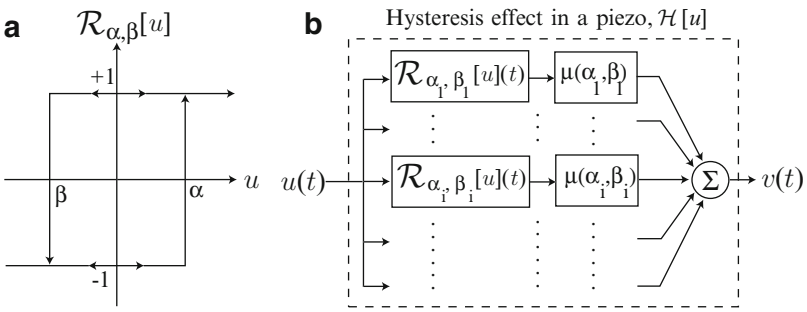


Fig. 7.10 (**a**) The elementary relay (Preisach hysteron) and (**b**) the output $v(t)$ is the sum of weighted hysterons

where $\mu(\alpha, \beta)$ is called the Preisach weighting function, and (α, β) belongs to the Preisach plane **P**, defined as,

$$\mathbf{P} \triangleq \{(\alpha, \beta) | \alpha \geq \beta; \underline{u} \leq \alpha, \beta \leq \bar{u}\}. \tag{7.13}$$

The Preisach hysteresis model can be obtained experimentally from measured output data, for instance, by applying an appropriate input voltage and measuring the piezoactuator's displacement response. In this case, based on the work of Banks et al. [55], a discrete form of the output Eq. (7.12) can be approximated by

$$v(t) \approx \sum_{i=1}^{N} R^* \mu_r A_r, \tag{7.14}$$

where A_i represents the area associated with the ith node, μ_i is the average value of the weighting surface over area A_i and R^* takes on value $+1$ or -1 depending on the state of the node, or relay at the node.

7.3.2.5 Prandlt-Ishlinksii Model

Another operator-type hysteresis model which has recently been investigated for piezoactuators is the Prandlt-Ishlinskii model, a phenomenological model that is characterized by a play operator or a stop operator [56, 57]. This operator is continuous and rate-independent, as shown in Eq. (7.15). If the input u is continuous and monotone in subspace $C[t_i, t_{i+n}]$ of $C[0, t]$, then the operator maps function $H_r : C[0, t] \rightarrow C[0, t]$, that is,

$$\begin{aligned} H_r[u](o) &= h_r(u(0), 0) = 0, \\ H_r[u](t) &= h_r(u(t), H_r[u](t)). \end{aligned} \tag{7.15}$$

In the operator function, $h_r(u(t), H_r[u](t_i)) = \max(u - r, \min(u + r, H_r[u](t_{i-1})))$. The square bracket of the operator $(H_r[\cdot])$ means that the operator is dependent on input function. The operator maps an input function to an output function, which represents the basic behavior of hysteresis, as illustrated in Fig. 7.11. The r in the $h_r(\cdot)$ is the threshold. The threshold r decides when the operator is turned on, therefore with different r the operator maps the input and output curve in different shapes as shown in Fig. 7.11.

With this operator the P-I model describes the input u and hysteresis output y_h by

$$y_h(t) = ku + \int_0^R p(r) H_r[u](t) dr, \tag{7.16}$$

where k is the positive gain, and $p(r)$ is the density function, which affects the shape and size of the hysteresis curve.

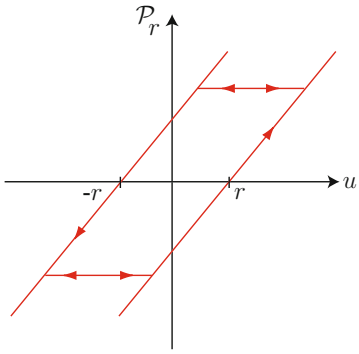


Fig. 7.11 The hysteresis play operator

7.4 Feedback Control for Nanopositioning

A feedback controller works by comparing the reference behavior trajectory (such as displacement) to the actual response (also displacement), see Fig. 7.1a. By minimizing the positioning error, a feedback controller can compensate for all forms of positioning errors that are within its effective bandwidth. Feedback control is the most commonly used technique for eliminating positioning errors in nanopositioning systems due to its simplicity and robustness. This section provides an overview of feedback control techniques, including high-gain feedback and damping-type controllers such integral resonant control. When the reference trajectory is periodic, repetitive control can significantly improve the tracking performance of a feedback loop. The repetitive control approach is briefly described for nanopositioning systems.

Feedback control schemes such as traditional PID [5], state-feedback [6], gain scheduling [58], and H_∞ control [3, 7] handle modeling errors and are robust with respect to parameter variation. Although feedback control has been applied extensively in nanopositioning, the low gain margin in piezo positioners tend to limit the performance, i.e., high feedback gain tends to destabilize such systems [5, 11]. In practice, a compromise is sought between performance and instability; feedback gains (such as proportional, derivative, and integral terms) are adjusted to improve performance without instability [12, 13]. Additionally, the precision of feedback controllers can be sensitive to sensor noise and can be bandwidth limited.

Feedback controllers such as standard proportional-integral (PI) controllers are commonly used due to their robustness and high low-frequency gain, but PI controllers offer limited closed-loop bandwidth. It was shown that for a nanopositioner with second-order dynamics, the maximum closed-loop bandwidth $\bar{\omega}_{BW}$ is [59]

$$\bar{\omega}_{BW} \leq 2\omega_n\zeta, \tag{7.17}$$

where ω_n is the natural frequency and ζ is the damping ratio of the positioning system. With this, low damping ratio severely limits the closed-loop bandwidth of a PI-controlled system.

The simplest and most popular approach to suppress the sharp resonance behavior is to use notch filters or inversion-based filters in the closed-loop [6, 25, 60]. By doing this, the gain margin and closed-loop bandwidth can be improved, even up to the resonance frequency [60]. For example, the notch filter of the following form can be used

$$D(s) = k_D \frac{(s - 2\pi z_1)(s - 2\pi z_2)}{(s - 2\pi p_1)(s - 2\pi p_2)}, \tag{7.18}$$

where k_D is a constant, z_1 and z_2 are the zeros of the filter that *cancel* the effect of the sharp resonant peak of the nanopositioning system, and p_1 and p_2 are poles to ensure a proper transfer function $D(s)$. Such notch filters can be realized using analog op-amp circuits (e.g., [61], pp. 394–399).

Figure 7.12 compares the measured frequency response for a piezoscanner, a notch filter, and the composite system. The measured gain margin of the original system is −17.05 dB, whereas the gain margin of the composite system is 30.86 dB. Thus, the notch filter significantly increase in the gain margin to enable high-gain feedback control to compensate for effects such as hysteresis and creep, as well as improve transient response [25]. It is pointed out that employing notch filters

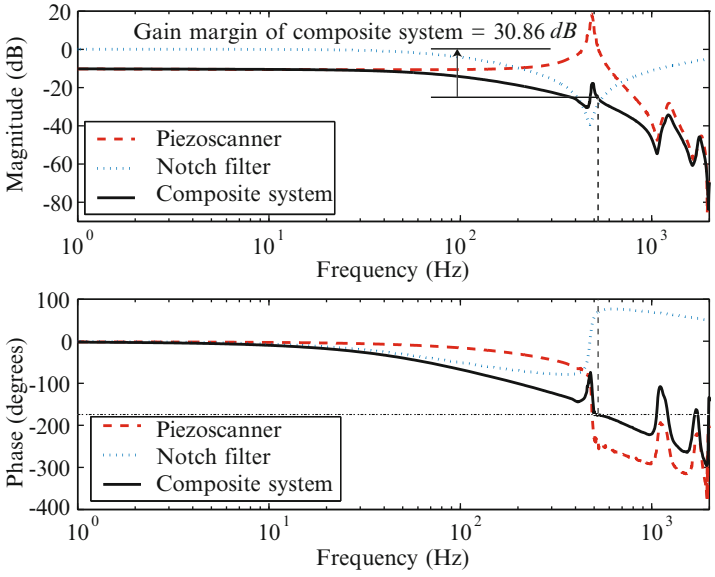


Fig. 7.12 Measured frequency response of the piezoscanner (*dashed line*), a notch filter (*dotted line*), and a notch filter cascaded with the piezoscanner (*solid line*). The measured gain margin of the original system is −17.05 dB, whereas the gain margin of the composite system is 30.86 dB

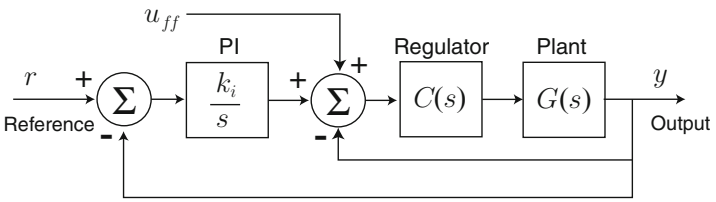


Fig. 7.13 Advanced integral resonant control (IRC) with integral action and a feedforward input u_{ff} [69]

requires knowledge of the resonance frequency. Even a small shift in the resonance behavior as low as 10 %, say caused by changes in the payload mass, can lead to instability of the closed-loop system.

Robust [8], H_∞-based [62, 63], and loop-shaping [64] control techniques have been considered for high-bandwidth operation. Other damping control techniques include positive position feedback (PPF) control and polynomial-based controller designs [65, 66]. In the PPF approach, an inner feedback loop is used to damp the highly resonant mode of the positioner. Application of the PPF controller on a commercial AFM system demonstrated damping as well as cross-coupling compensation for a piezoelectric tube scanner [67]. Integral resonant control (IRC), a method introduced as a means for augmenting the structural damping of resonant systems with collocated sensors and actuators [68], damps the vibration and makes the system robust to the unmodeled dynamics and resonance frequency variations due to changes in the payload [69]. The block diagram of an advanced version of the IRC approach which includes integral action and a feedforward input u_{ff} is illustrated in Fig. 7.13, where $G(s)$ is the plant and the regulator $C(s)$ is

$$C(s) = \frac{-k}{s - kD_f}, \tag{7.19}$$

k is the constant, and D_f is the feedthrough term [69]. Other active damping techniques include receding horizon control [70, 71] as well as active and passive shunt damping techniques for piezoelectric tube scanners [72, 73].

Recently, a feedback-based approach known as repetitive control (RC) was exploited for tracking periodic trajectories in nanopositioning applications including high-speed and metrological AFMs [74–76]. The RC provides high gain at the harmonics of the reference trajectory by incorporating a signal generator within the feedback loop [77, 78]. For piezo-based nanopositioning systems, the RC can be designed for low tracking error in the presence of dynamic and hysteresis effects [79, 80]. Some advantages of this approach include minimal system modeling, robustness due to the feedback structure, and straightforward digital implementation; however, *a priori* knowledge of the period of the reference trajectory is required.

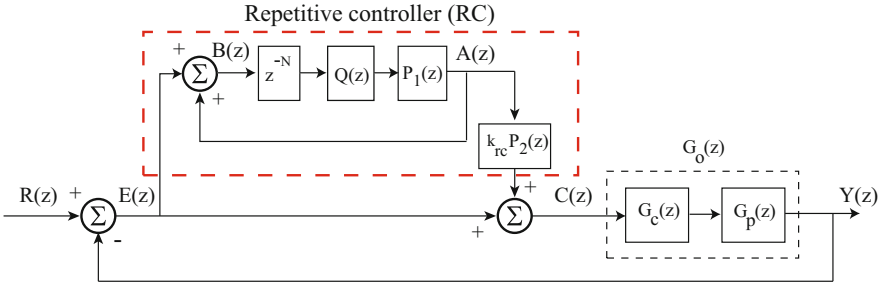


Fig. 7.14 The repetitive control (RC) feedback system

To illustrate the concept of RC, consider an example of the discrete-time RC closed-loop system shown in Fig. 7.14. The dynamics of the nanopositioner, assumed to be linear, are represented by $G_p(z)$, where $z = e^{j\omega T_s}$, $\omega \in (0, \pi/T_s)$. In the block diagram, $G_c(z)$ is a feedback controller, such as an existing PID controller; $Q(z)$ is a low-pass filter for robustness; k_{rc} is the RC gain; and $P_1(z) = z^{m_1}$ and $P_2(z) = z^{m_2}$, where m_1, m_2 are non-negative integers, are positive phase lead compensators to enhance the performance of the RC feedback system. It is emphasized that the phase lead compensators z^{m_1} and z^{m_2} provide a linear phase lead of (in units of radians)

$$\theta_{1,2}(\omega) = m_{1,2} T_s \omega, \text{ for } \omega \in (0, \pi/T_s). \tag{7.20}$$

The key component of the repetitive controller is the signal generator. To create a signal generator with period T_p, the inner loop contains the pure delay z^{-N}, where the positive integer $N = T_p/T_s$ is the number of points per period T_p; and T_s is the sampling time. An analysis of the performance of the closed-loop system is presented below, where the following assumptions are considered: (1) the reference trajectory $R(z)$ is periodic and has period T_p and (2) the closed-loop system without the RC loop is asymptotically stable, i.e., $1 + G_c(z)G_p(z) = 0$ has no roots outside of the unit circle in the z-plane.

Assumptions 1 and 2 are easily met for many applications in nanopositioning, including SPMs. For example in AFM imaging the lateral movements of the piezoactuator are periodic, such as a triangle scanning signal. Also, most SPMs are equipped with feedback controllers $G_c(z)$ to control the lateral positioning, which can be tuned to be stable.

The transfer function of the signal generator (or RC block, Fig. 7.14) that relates $E(z)$ to $A(z)$ is given by

$$\frac{A(z)}{E(z)} = \frac{Q(z)P_1(z)z^{-N}}{1 - Q(z)P_1(z)z^{-N}} = \frac{Q(z)z^{(-N+m_1)}}{1 - Q(z)z^{(-N+m_1)}}. \tag{7.21}$$

In the absence of both the low-pass filter $Q(z)$ and positive phase lead $P_1(z) = z^{m_1}$, the poles of the signal generator are $1 - z^{-N} = 0$; therefore, the frequency response of the signal generator reveals infinite gain at the fundamental frequency and its harmonics $\omega = 2n\pi/T_p$, where $n = 1, 2, 3, \ldots$. The infinite gain at the harmonics is what gives the RC its ability to track a periodic reference trajectory. As a result, RC is a useful control method for applications such as SPM in which the scanning motion is repetitive. Unfortunately, the RC also contributes phase lag which causes instability. Therefore, the stability, robustness, and tracking performance of the RC closed-loop system must be carefully considered.

For the RC system shown in Fig. 7.14, it has been shown that for stability the following two conditions for the RC gain k_{rc} and linear phase lead $\theta_2(\omega)$ must be met [74]:

$$0 < k_{rc} < \frac{2\cos[\theta_T(\omega) + \theta_2(\omega)]}{A(\omega)} \quad \text{and} \tag{7.22}$$

$$-\pi/2 < [\theta_T(\omega) + \theta_2(\omega)] < \pi/2. \tag{7.23}$$

In Eq. (7.23), the lead compensator $P_2(z) = z^{m_2}$ accounts for the phase lag of the closed-loop feedback system without RC. Enhanced designs of RC for nanopositioning include dual-stage RC [79] and RC that accounts for hysteresis behavior [80].

The tracking results comparing PID, regular RC, and the RC with the phase lead compensators for ±25-μm scanning at 5, 10, and 25 Hz are presented in Fig. 7.15 and Table 7.1. The steady-state tracking errors, measured at the last two cycles, are reported as a percentage of the range of motion (maximum error and root-mean-squared error). Because the action of the repetitive controller is delayed by one scan period, the tracking response for the first period are similar for the PID, RC, and RC with phase lead compensation as shown in Fig. 7.15. However, after the first period the RC begins to take action as illustrated by the reduction of the tracking error from one cycle to the next. On the other hand, the tracking error of the PID controller persists from one cycle to the next.

The 5 Hz scanning results shown in Fig. 7.15(a1) and (b1) and Table 7.1 demonstrate that the regular RC controller reduced maximum tracking error from 2.01 to 0.96 % compared to the PID controller, a 52 % reduction. By using RC with the phase lead compensation, an additional 55 % improvement in tracking performance was achieved. In this case, the maximum tracking error is 0.43 %.

Table 7.1 Tracking results for ±25-μm range

	5 Hz		10 Hz		25 Hz	
Controller	$e_{\max}(\%)$	$e_{\text{rms}}(\%)$	$e_{\max}(\%)$	$e_{\text{rms}}(\%)$	$e_{\max}(\%)$	$e_{\text{rms}}(\%)$
PID	2.01	1.28	3.99	2.61	9.16	6.61
RC	0.96	0.21	2.74	0.79	8.86	3.69
RC + phase leads	0.43	0.08	0.46	0.10	1.78	0.57

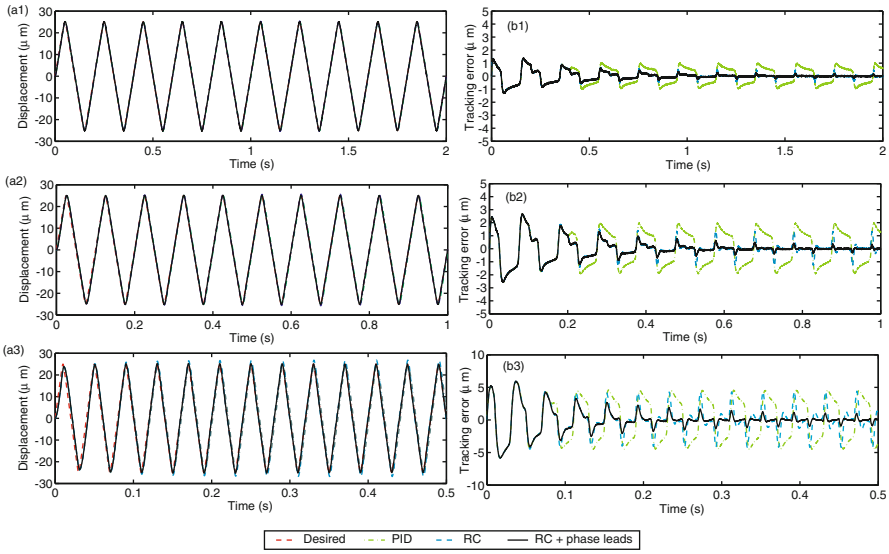


Fig. 7.15 Experimental tracking response and error for PID (*dash-dot*), RC (*dashed line*), and RC with phase lead compensation [$m_1 = 6$ and $m_2 = 0$] (*solid line*) for 5 Hz (*a1* and *b1*), 10 Hz (*a2* and *b2*), and 25 Hz (*a3* and *b3*) scanning

At 25 Hz, the tracking error of PID was unacceptable large at 9.16 %. In fact, for AFM scanning operations the maximum tracking error should be less than a few percent. The results in Table 7.1 show that the regular plug-in RC controller was not able to improve the tracking performance at 25 Hz. However, the RC with phase lead compensation gave lower maximum tracking error at 1.78 %. Therefore, the RC with phase lead compensation enables precision tracking at higher scan rates.

Finally, scanning offset from the piezoactuator's center position is demonstrated as shown in Fig. 7.16. For this offset scanning operation, the PID controller accounted for the low-frequency dynamics such as creep and the RC was used for tracking the periodic trajectory. The tracking results in Fig. 7.16 show that the RC was effective at minimizing the tracking error.

7.5 Feedforward Control for Nanopositioning Systems

Feedback controllers react to the measured tracking error; however, feedforward controllers compensate or anticipate for deficit performance. As the operating speed of a piezo-based nanopositioner increases, performance is limited by essentially dynamic and hysteresis effects [1, 81]. Feedforward control is an open-loop approach as depicted in Fig. 7.1b. As shown, an inverse model produces the feedforward input u_{ff} that is applied to the positioning system. The accuracy of

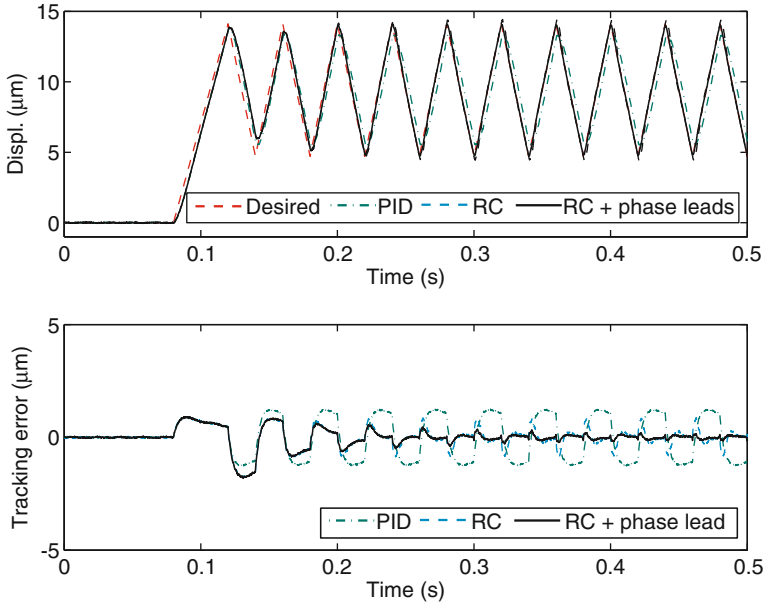


Fig. 7.16 Tracking results for offset triangle scan at 25 Hz

feedforward control, for example how close the actual output y matches the desired output y_d, depends on the *quality* of the inverse model and whether external disturbances are present. Being an open-loop approach, feedforward control is subject to certain shortcomings, namely lack of robustness. However, for applications such as vibration compensation with a reasonably accurate model of the system dynamics, the advantages of feedforward control outweighs its disadvantages. In particular, feedforward control can provide high-bandwidth positioning, exceeding that of feedback-based methods. Also, feedforward control does not require continuous sensor feedback, and thus sensor-noise related issues can be avoided entirely.

To improve robustness, feedforward can be integrated with feedback control, as well to account for nonlinearity such as hysteresis [25]. Integrated feedback and feedforward control also eliminates the need for modeling and inverting nonlinear behaviors. Such a task can be difficult and computationally demanding. If iteration is allowed, iterative feedforward techniques as discussed below provides good performance with minimal modeling.

7.5.1 Simple DC-Gain Feedforward Control

At frequencies well below the dominant resonant peak, a simple feedforward input u_{ff} can be computed by scaling the desired output trajectory $y_d(t)$ by the inverse of the DC-gain $G(0)$ of the system $G(s)$. For example,

$$u_{ff}(t) = \frac{1}{G(0)} y_d(t). \tag{7.24}$$

Simple DC-gain feedforward is often used when a plant model is difficult and/or expensive to obtain. Also, if the frequency components of the desired trajectory are well below the dominant resonances, the need for inverting high-frequency dynamics for feedforward is less important. It is noted that simple DC-gain feedforward offers good performance at low-frequency and over small displacement ranges, and also under the assumption that the DG-gain remains stable.

7.5.2 *Inversion-Based Feedforward Control for Linear Dynamics*

At higher frequencies, the effects of dynamics must be considered in the feedforward control input. In this case, the problem becomes inverting a dynamics model $G(s)$ to find the feedforward input $u_{ff}(t)$ over a specific frequency range. For piezoactuators, the dynamics include vibration and the creep effect [30].

Let $G(s)$ be a transfer function model that captures the linear dynamics of the piezoactuator. In the frequency domain, the feedforward input that accounts for the dynamics is given by

$$u_{ff}(j\omega) = G^{-1}(j\omega) y_d(j\omega), \tag{7.25}$$

where $G^{-1}(j\omega)$ is the inverse dynamics model.

In Eq. (7.25), the Fourier transform of the desired output trajectory $y_d(j\omega)$ and a plant model $G(j\omega)$ are needed to determine the feedforward input. If a measured frequency response function $G(j\omega)$ is available, for example measured by a dynamic signal analyzer, then the experimental data can be used directly to compute the feedforward input $u_{ff}(j\omega)$. Then, the time-domain solution to Eq. (7.25) is found by taking the inverse Fourier transform of $u_{ff}(j\omega)$.

One key feature of this feedforward technique is that it can be applied to nonminimum-phase systems [82–84]. Although the dynamic effects are specifically addressed in this section, the approach can be combined with alternative feedforward or feedback methods that compensate for hysteresis when the range of motion becomes large [25].

The inversion-based method presented above may yield excessively large inputs when the system has lightly damped system zeros. These large inputs can saturate the voltage amplifiers that drive the piezoactuator, or, even worse, depole the piezoactuator. Additionally, large model uncertainties around the resonant peaks or lightly damped zeros can cause significant error in computing the feedforward input. These model uncertainties thus lead to a lack of robustness when the inversion-based feedforward method is used. The following optimal inversion approach is used to

account for these issues. Specifically, an optimal feedforward input is obtained by minimizing the quadratic cost function [85]

$$J(u) = \int_{-\infty}^{\infty} \left\{ u^*(j\omega) R(j\omega) u(j\omega) + [x(j\omega) - x_d(j\omega)]^* \, Q(j\omega) \, [x(j\omega) - x_d(j\omega)] \right\} d\omega, \tag{7.26}$$

where '*' denotes the conjugate transpose, and $R(j\omega)$ and $Q(j\omega)$ are non-negative, frequency-dependent real-valued weights on the input energy and the tracking error, respectively. The optimal feedforward input $u_{ff,opt}$ that minimizes (7.26) is

$$u_{ff,opt}(j\omega) = \left[\frac{G^*(j\omega) Q(j\omega)}{R(j\omega) + G^*(j\omega) Q(j\omega) G(j\omega)} \right] y_d(j\omega). \tag{7.27}$$

By choosing the frequency-dependent weights $R(j\omega)$ and $Q(j\omega)$, it is possible to systematically consider the effects of the input magnitude and the model uncertainties. For instance, the input energy weight $R(j\omega)$ can be chosen to be much larger than the tracking error weight $Q(j\omega)$ at frequencies where large model uncertainties exist or around lightly damped zeros. For details and implementation issues, see [86]. Other feedforward methods include input shaping techniques [87] and shaped-triangle technique for high-speed AFM [88].

Both linear dynamics and hysteresis behaviors can be compensated for using feedforward control by inverting the plant dynamics and nonlinearities (see Fig. 7.17) [30]. Both the dynamic and hysteresis effect are often intertwined [89], but can be modeled for feedforward control using the cascade structured shown in Fig. 7.17. Specifically, the range-dependent hysteresis effect is treated as a rate-independent, input nonlinearity represented by H. The vibrational dynamics (and creep effect) are captured by the linear dynamics model $G(s)$. The cascade model structure is used extensively to model piezoactuators and similar systems [30, 47].

It is pointed out that the model-based feedforward approach cannot correct for positioning errors due to plant uncertainties and external perturbations [90]; but feedback control can be used to minimize the uncertainties and then combined with feedforward control [25, 91–94].

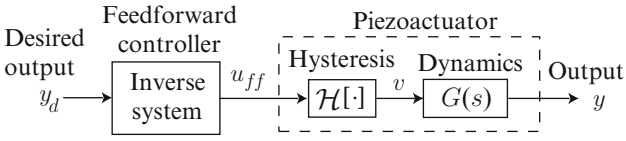


Fig. 7.17 The feedforward control approach for piezoactuator system

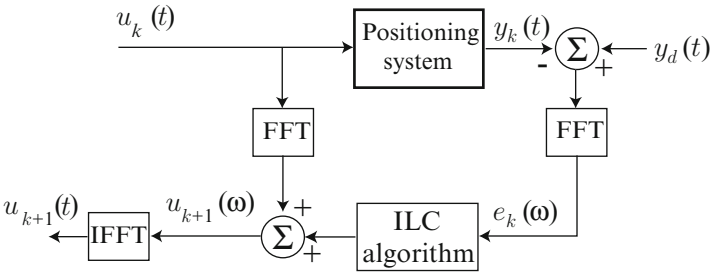


Fig. 7.18 Block diagram of frequency-domain ILC scheme

7.5.3 Iterative Feedforward Control

Many nanopositioning applications, for example SPM imaging, require tracking a repeating (periodic) trajectory. The scanning motion in this case repeats from one cycle to the next. Rather than model and invert the dynamics and nonlinearities of a positioning system for feedforward control, for repeating scanning motion, the feedforward input can be found using iterative techniques for both dynamics and hysteresis compensation [39, 95, 96]. This approach is commonly referred to as iterative learning control (ILC), where recently an approximate inverse of the plant dynamics G was used in the input-update law [95]

$$u_{k+1}(\omega) = u_k(\omega) + \rho(\omega)\hat{G}^{-1}(\omega)\big[y_d(\omega) - y_k(\omega)\big], \tag{7.28}$$

to achieve faster convergence (k is the iteration number). A block diagram of the frequency-domain implementation of the ILC control law (7.28) is shown in Fig. 7.18. Particularly, the input $u_k(t)$ and tracking error $e_k(t)$ are Fourier transformed, then the ILC algorithm is applied, producing the updated input $u_{k+1}(\omega)$. The time-domain input $u_{k+1}(t)$ is obtained by inverse Fourier transform. Then the input is applied to the system and the process is repeated [28, 97]. A model-less ILC approach where the inverse G^{-1} is obtained from the measured input-output data from the previously iteration step was considered for AFM applications [98–100].

Piezo-based nanopositioners exhibit nonlinearity such as hysteresis. Although the ILC approach is well suited for precise control of linear dynamics, the challenge is developing a convergence criteria for hysteresis. The difficulty in proving convergence of ILC algorithms for hysteretic systems arises due to two main reasons: (1) branching effects and (2) nonlinearity of each branch [101].

The inability, to predict the direction in which the input needs to be changed for reducing the output error in hysteretic systems, can be overcome if the input-output behavior is restricted to belong on one single hysteresis branch. This observation that the direction can be determined from the output error on a single branch was used to prove the convergence for an ILC algorithm for hysteretic systems in [39]. First, the desired output trajectory is partitioned into monotonic sections (several

branches). The algorithm is applied to each section until a desired tracking precision is achieved.

In particular, the ILC algorithm (ILCA) of the following form can be used to compensate for hysteresis [39]:

$$u_{k+1}(t) = u_k(t) + \rho[y_d(t) - y_k(t)], \forall t \in I. \tag{7.29}$$

The ILCA Equation (7.29) converges if the desired trajectory $v_d(t)$ is continuous and monotonic over the finite time interval I. The iteration gain ρ for convergence is based on the parameters of the hysteresis model, such as the Preisach hysteresis model. It is pointed out that a sufficiently small iteration gain can be found provided the output different can be bounded above and below by the input difference [39].

7.5.4 Adding Feedback to Feedforward Control

Modeling and inverting the dynamic and hysteresis effects are effective methods for precision positioning in piezoactuators in nanopositioning systems [30, 86]. However, because the approach exploits knowledge of the piezoactuator behavior, the modeling process can be time-consuming, particularly when both the inverse dynamics and inverse hysteresis are used. If a simpler method to account for hysteresis is preferred over control performance, then high-gain feedback control can be used to linearize the nonlinear behavior of the piezoactuator [25]. In this case, the vibrational dynamics are modeled and inverted as described above, then the feedforward input is combined with the feedback controller as shown in Fig. 7.1c. However, when feedback is used, piezoactuators often exhibit low gain margin and can cause instability. Thus, as mentioned earlier, notch filters can be used to suppress the dominant resonant peaks.

7.5.5 Experimental Tracking Results

The results shown in Fig. 7.19 and Table 7.2 demonstrate the performance of feedforward control and integrated feedforward and feedback control applied to a nanopositioning system in an AFM. A high-gain feedback controller as described in [25] was used to compensate for hysteresis and creep. The optimal inversion feedforward controller was used to minimize the vibrational dynamics. The scan range for the experiment is $50.00\,\mu m$. The weightings of the optimal inversion controller were chosen to give up tracking of all frequency components beyond $450\,Hz$ (to avoid saturation). In particular, the weightings are $Q = 1$ and $R = 0$ for $\omega \leq 450\,Hz$, and $Q = 0$ and $R = 1$ for $\omega > 450\,Hz$.

As shown by the results, the integrated feedback and optimal inversion feedforward approach substantially reduces positioning errors compared to just using

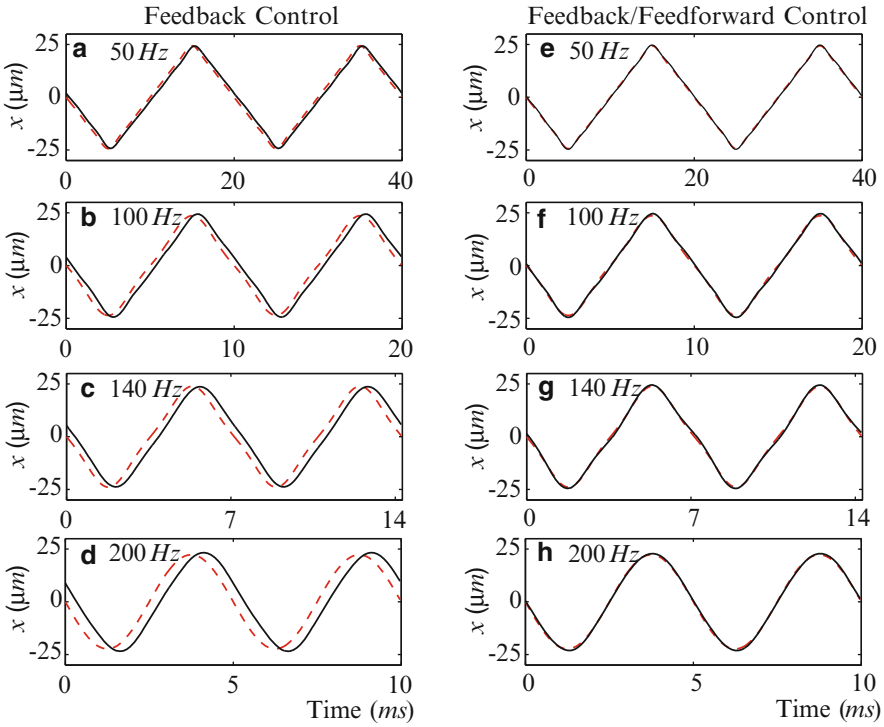


Fig. 7.19 Experimental results: tracking of triangular reference trajectory. (**a**)–(**d**) Feedback control. (**e**)–(**h**) Feedback and optimal inversion feedforward control. Optimal weightings were chosen as $Q = 1$ and $R = 0$ for $\omega \leq 450\,Hz$, and $Q = 0$ and $R = 1$ for $\omega > 450\,Hz$. *Solid line* is the measured response and *dotted line* is the desired trajectory

Table 7.2 Tracking performance for triangular reference trajectory using integrated feedback and optimal inversion feedforward control

	Feedback		Integrated feedback/feedforward	
Scan rate (Hz)	e_{max} (%)	e_{rms} (%)	e_{max} (%)	e_{rms} (%)
50	4.37	3.58	1.54	0.74
100	9.49	6.79	2.95	1.39
140	12.11	9.33	2.94	1.42
200	18.15	13.69	2.03	1.18

Values reported as percentage of total output range (50.00 μm)

feedback control for tracking a triangular trajectory as illustrated in Fig. 7.19 and Table 7.2. Good tracking (i.e., $e_{max} \leq 5\,\%$) is achieved even at the 200 Hz scan rate. Furthermore, the maximum and root-mean-square of the tracking error were reduced by over 88 and 91 % at this scan rate compared to using only feedback control.

7.6 Cross-Coupling Effects, Charge Control, and Thermal Drift

Careful mechanical design can significantly minimize cross-coupling effects [76]. However, during high-speed operation such as fast raster scanning in video-rate AFMs, the cross-coupling effect cannot be ignored as inertial effects and poor boundary conditions can cause out-of-plane modes to be excited [27]. Control methods to deal with cross-coupling issues include inverse-based feedforward control [28], H_∞-based control [27, 102, 103], iterative feedforward approach [104], and adaptive control [105].

By controlling the charge (rather than voltage) applied to the piezoactuator, hysteresis can be reduced. Early demonstration of the effectiveness of charge drives have appeared in the literature [106–109]. More recently, advancements in the design of charge drives for piezoelectric actuators have resulted in effective designs for operating at low and high frequencies [110–112]. These charge drives have also been integrated with inversion-based feedforward control for piezoelectric actuators [113] and SPM [112].

Thermal expansion can severely limit the precision, repeatability, and overall performance of a nanopositioning system. Specifically, the result includes temperature-dependent drift in the motion and high thermal stresses which ultimately lead to cracking, warping, or loosening of components. Unfortunately, for positioning at the nanometer scale, thermal effects cannot be ignored. In a typical AFM operated in ambient temperature, even a small 1° change in temperature can cause as much as 50 nm drift [114]. Feedback controllers have been used successfully to minimize thermal drift [115]. For SPMs, an effective method to deal with drift is to measure and compensating for the effect through consecutive SPM images [116–119].

7.7 Summary

This chapter focused on feedback and feedforward controllers for nanopositioning systems. First, the effects of vibration, hysteresis, and creep were described. Simple methods traditionally employed to avoid these effects were discussed. Afterwards, the modeling of behavior of nanopositioners was presented and these models are typically used in control system design and simulation. Finally, popular feedback and feedforward controllers to handle vibration, hysteresis, and creep were presented. The controllers included IRC, repetitive control, model-based inversion feedforward control, and iterative feedforward control methods.

References

1. G.M. Clayton, S. Tien, K.K. Leang, Q. Zou, S. Devasia, A review of feedforward control approaches in nanopositioning for high speed SPM. ASME J. Dyn. Syst. Meas. Control. **131**(6), 061101 (19 pages) (2009)
2. S. Devasia, E. Eleftheriou, S.O.R. Moheimani, A survey of control issues in nanopositioning. IEEE Trans. Control Syst. Technol. **15**(5), 802–823 (2007)
3. S.M. Salapaka, M.V. Salapaka, Scanning probe microscopy. IEEE Control. Syst. Mag. **28**(2), 65–83 (2008)
4. A.J. Fleming, K.K. Leang, *Design, Modeling and Control of Nanopositioning Systems* (Springer, New York, 2014)
5. R.C. Barrett, C.F. Quate, Optical scan-correction system applied to atomic force microscopy. Rev. Sci. Instrum. **62**(6), 1393–1399 (1991)
6. Y. Okazaki, A micro-positioning tool post using a piezoelectric actuator for diamond turning machines. Precis. Eng. **12**(3), 151–156 (1990)
7. S. Korson, A.J. Helmicki, An h_∞ based controller for a gas turbine clearance control system, in *Proceedings of the IEEE Conference on Control Applications* (1995), pp. 1154–1159
8. S. Salapaka, A. Sebastin, J.P. Cleveland, M.V. Salapaka, High bandwidth nano-positioner: a robust control approach. Rev. Sci. Instrum. **73**(9), 3232–3241 (2002)
9. F. Lowrie, M. Cain, M. Stewart, M. Gee, Time dependent behaviour of piezo-electric materials. Technical Report 151, National Physical Laboratory, 1999
10. H.-J. Lee, D.A. Saravanos, The effect of temperature dependent material properties on the response of piezoelectric composite materials. J. Intell. Mater. Syst. Struct. **9**(7), 503–508 (1998)
11. J.A. Main, E. Garcia, Piezoelectric stack actuators and control system design: strategies and pitfalls. AIAA J. Guid. Control. Dyn. **20**(3), 479–485 (1997)
12. F. Braet, R.D. Zanger, S. Kammer, E. Wisse, Noncontact versus contact imaging: an atomic force microscope study on hepatic endothelial cells in vitro. Int. J. Imaging Syst. Technol. **8**(2), 162–167 (1997)
13. G. Schitter, P. Menold, H.F. Knapp, F. Allgöwer, A. Stemmer, High performance feedback for fast scanning atomic force microscopes. Rev. Sci. Instrum. **72**(8), 3320–3327 (2001)
14. J. Schafer, H. Janocha, Compensation of hysteresis in solid-state actuators. Sensors Actuators A **49**, 97–102 (1995)
15. P. Ge, M. Jouaneh, Tracking control of a piezoceramic actuator. IEEE Trans. Control Syst. Technol. **4**(3), 209–216 (1996)
16. H. Janocha, K. Kuhnen, Real-time compensation of hysteresis and creep in piezoelectric actuators. Sensors Actuators A **79**, 83–89 (2000)
17. S. Mittal, C.-H. Menq, Hysteresis compensation in electromagnetic actuators through preisach model inversion. IEEE/ASME Trans. Mechatro. **5**(4), 394–409 (2000)
18. S. Majima, K. Kodama, T. Hasegawa, Modeling of shape memory alloy actuator and tracking control system with the model. IEEE Trans. Control Syst. Technol. **9**(1), 54–59 (2001)
19. O.M.E. Rifai, K. Youcef-Toumi, Creep in piezoelectric scanners of atomic force microscopes, in *Proceeding of the American Control Conference* (2002), pp. 3777–3782
20. R.C. Smith, M.V. Salapaka, A. Hatch, J. Smith, T. De, Model development and inverse compensator design for high speed nanopositioning, in *Proceeding of the 41st IEEE Conference on Decision and Control* (2002), pp. 3652–3657
21. D. Croft, D. McAllister, S. Devasia, High-speed scanning of piezo-probes for nano-fabrication. ASME J. Manuf. Sci. Eng. **120**(3), 617–622 (1998)
22. D. Croft, S. Devasia, Vibration compensation for high speed scanning tunneling microscopy. Rev. Sci. Instrum. **70**(12), 4600–4605 (1999)
23. S. Stilson, A. McClellan, S. Devasia, High-speed solution switching using piezo-based micro-positioning stages. in *Proceedings of the American Control Conference* (2001),pp. 2238–2243

24. S. Devasia, Should model-based inverse inputs be used as feedforward under plant uncertainty? IEEE Trans. Autom. Control **47**(11), 1865–1871 (2002)
25. K.K. Leang, S. Devasia, Feedback-linearized inverse feedforward for creep, hysteresis, and vibration compensation in AFM piezoactuators. IEEE Trans. Control Syst. Technol. **15**(5), 927–935 (2007)
26. T.R. Hicks, P.D. Atherton, *The Nanopositioning Book* (Queensgate Instruments Ltd., Torquay, 1997)
27. Y. Yong, S.O.R. Moheimani, B.J. Kenton, K.K. Leang, Invited review: High-speed flexure-guided nanopositioning: mechanical design and control issues. Rev. Sci. Instrum. **83**(12), 121101 (2012)
28. S. Tien, Q. Zou, S. Devasia, Iterative control of dynamics-coupling-caused errors in piezoscanners during high-speed AFM operation. IEEE Trans. Control Syst. Technol. **13**(6), 921–931 (2005)
29. A.E. Holman, P.M.L.O. Scholte, W.C. Heerens, F. Tuinstra, Analysis of piezo actuators in translation construction. Rev. Sci. Instrum. **66**(5), 3208–3215 (1995)
30. D. Croft, G. Shed, S. Devasia, Creep, hysteresis, and vibration compensation for piezoactuators: atomic force microscopy application. ASME J. Dyn. Syst. Meas. Control. **123**(1), 35–43 (2001)
31. T. Sulchek, R. Hsieh, J.D. Adams, S.C. Minne, C.F. Quate, High-speed atomic force microscopy in liquid. Rev. Sci. Instrum. **71**(5), 2097–2099 (2000)
32. G. Schitter, K.J. Astrom, B.E. DeMartini, P.J. Thurner, K.L. Turner, P.K. Hansma, Design and modeling of a high-speed AFM-scanner. IEEE Trans. Control Syst. Technol. **15**(5), 906–915 (2007)
33. K.K. Leang, A.J. Fleming, High-speed serial-kinematic AFM scanner: design and drive considerations, in *American Control Conference, Invited Session on Modeling and Control of SPM* (2008), pp. 3188–3193
34. T. Ando, N. Kodera, D. Maruyama, E. Takai, K. Saito, A. Toda, A high-speed atomic force microscope for studying biological macromolecules in action. Jpn. J. Appl. Phys., Part 1 **41**(7B), 4851–4856 (2002)
35. D.C. Jiles, D.L. Atherton, Theory of ferromagnetic hysteresis. J. Magn. Magn. Mater. **61**, 48–60 (1986)
36. R. Cross, *Unemployment, Hysteresis, and the Natural Rate Hypothesis* (Basil Blackwell Ltd., New York, 1988)
37. H. Cao, A.G. Evans, Nonlinear deformation of ferroelectric ceramics. J. Am. Ceram. Soc. **76**, 890–896 (1993)
38. R.S. Robinson, Interactive computer correction of piezoelectric creep in scanning tunneling microscopy images. J. Comput.-Assist. Microsc. **2**(1), 53–58 (1996)
39. K.K. Leang, S. Devasia, Design of hysteresis-compensating iterative learning control for piezo positioners: application to atomic force microscopes. Mechatronics **16**(3–4), 141–158 (2006)
40. S.M. Hues, C.F. Draper, K.P. Lee, R.J. Colton, Effect of PZT and PMN actuator hysteresis and creep on nanoindentation measurements using force microscopy. Rev. Sci. Instrum. **65**(5), 1561–1565 (1994)
41. K.R. Koops, P.M.L.O. Scholte, W.L. de Koning, Observation of zero creep in piezoelectric actuators. Appl. Phys. A **68**, 691–697 (1999)
42. H. Jung, D.-G. Gweon, Creep characteristics of piezoelectric actuators. Rev. Sci. Instrum. **71**(4), 1896–1900 (2000)
43. W.D. Callister, *Materials Science and Engineering: An Introduction* (Wiley, New York, 1994)
44. G. Schitter, A. Stemmer, Fast closed loop control of piezoelectric transducers. J. Vac. Sci. Technol. B **20**(1), 350–352 (2002)
45. H. Jung, J.Y. Shim, D Gweon, New open-loop actuating method of piezoelectric actuators for removing hysteresis and creep. Rev. Sci. Instrum. **71**(9), 3436–3440 (2000)
46. P. Krejci, K. Kuhnen, Inverse control of systems with hysteresi and creep. IEE Proc. Control Theory Appl. **148**(3), 185–192 (2001)

47. X. Tan, J.S. Baras, Adaptive identification and control of hysteresis in smart materials. IEEE Trans. Autom. Control **50**(6), 827–839 (2005)
48. L.E. Malvern, *Introduction to the Mechanics of a Continuous Medium* (Prentice-Hall, Englewood Cliffs, 1969)
49. M. Hu, H. Du, S.-F. Ling, Z. Zhou, Y. Li, Motion control of an electrostrictive actuator. Mechatronics **14**(2), 153–161 (2004)
50. M. Goldfarb, N. Celanovic, Modeling piezoelectric stack actuators for control of micromanipulation. IEEE Control. Syst. Mag. **17**(3), 69–79 (1997)
51. B.D. Coleman, M.L. Hodgdon, A constitutive relation for rate-independent hysteresis in ferromagnetically soft materials. Int. J. Eng. Sci. **24**(6), 897–919 (1986)
52. F. Preisach, Uber die magnetische nachwirkung. Z. Phys. **94**, 277–302 (1935)
53. P. Ge, M. Jouaneh, Modeling hysteresis in piezoceramic actuators. Precis. Eng. **17**(3), 211–221 (1995)
54. I.D. Mayergoyz, *Mathematical Models of Hysteresis* (Springer, New York, 1991)
55. H.T. Banks, A.J. Kurdila, G. Webb, Identification of hysteretic confluence operators representing smart actuators: convergent approximations. Technical Report, North Carolina State University CRSC, April 1997
56. M.A. Janaideh, C.-Y. Su, S. Rakheja, Development of the rate-dependent Prandtl-Ishlinskii model for smart actuators. Smart Mater. Struct. **17**, 035026 (11pp) (2008)
57. K. Kuhnen, Modeling, identification and compensation of complex hysteretic nonlinearities: a modified prandtl-ishlinskii approach. Eur. J. Control. **9**(4), 407–418 (2003)
58. R.J.E. Merry, N.C.T. de Kleijn, M.J.G. van de Molengraft, M. Steinbuch, Using a walking piezo actuator to drive and control a high precision stage. IEEE/ASME Trans. Mechatro. **14**(1), 21–31 (2009)
59. A.J. Fleming, Nanopositioning system with force feedback for high-performance tracking and vibration control. IEEE Trans. Mechatro. **15**(3), 433–447 (2010)
60. D.Y. Abramovitch, S. Hoen, R. Workman, Semi-automatic tuning of PID gains for atomic force microscopes, in *American Control Conference* (2008), pp. 2684–2689
61. H.Y-F. Lam, *Analog and Digital Filters: Design and Realization* (Prentice-Hall, New York, 1979)
62. A. Sebastian, S.M. Salapaka, Design methodologies for robust nanopositioning. IEEE Trans. Control Syst. Technol. **13**(6), 868–876 (2005)
63. A. Stemmer, G. Schitter, J.M. Rieber, F. Allgöwer, Control strategies towards faster quantitative imaging in atomic force microscopy. *Eur. J. Control.* **11**(4–5), 384–395 (2005)
64. A. Daniele, S. Salapaka, M.V. Salapaka, M. Dahleh, Piezoelectric scanners for atomic force microscopes: design of lateral sensors, identification and control, in *Proceeding of the American Control Conference* (1999), pp. 253–257
65. B. Bhikkaji, M. Ratnam, A.J. Fleming, S.O.R. Moheimani, High-performance control of piezoelectric tube scanners. IEEE Trans. Control Sys. Technol. **15**(5), 853–866 (2007)
66. B. Bhikkaji, M. Ratnam, S.O.R. Moheimani, PVPF control of piezoelectric tube scanners. Sensors Actuators A Phys. **135**(2), 700–712 (2007)
67. I.A. Mahmood, S.O.R. Moheimani, Making a commercial AFM more accurate and faster using positive position feedback control. Rev. Sci. Instrum. **80**(6), 063705 (8 pages) (2009)
68. S.S. Aphale, A.J. Fleming, S.O.R. Moheimani, Integral resonant control of collocated smart structures. Smart Mater. Struct. **16**, 439–446 (2007)
69. A.J. Fleming, S.S. Aphale, S.O.R. Moheimani, A new method for robust damping and tracking control of scanning probe microscope positioning stages. IEEE Trans. Nanotechnol. **9**(4), 438–448 (2010)
70. N. Kodera, H. Yamashita, T. Ando, Active damping of the scanner for high-speed atomic force microscopy. Rev. Sci. Instrum. **76**(5), 05378–1–053708–5 (2005)
71. A.J. Fleming, A.G. Wills, S.O.R. Moheimani, Sensor fusion for improved control of piezoelectric tube scanners. IEEE Trans. Control Sys. Technol. **15**(6), 1265–6536 (2008)
72. A.J. Fleming, S.O.R. Moheimani, Sensorless vibration suppression and scan compensation for piezoelectric tube nanopositioners. IEEE Control Sys. Technol. **14**(1), 33–44 (2006)

73. S.S. Aphale, A.J. Fleming, S.O.R. Moheimani, High speed nano-scale positioning using a piezoelectric tube actuator with active shunt control. Micro Nano Lett. **2**(1), 9–12 (2007)
74. U. Aridogan, Y. Shan, K.K. Leang, Design and analysis of discrete-time repetitive control for scanning probe microscopes. ASME J. Dyn. Syst. Meas. Control. **131**, 061103 (12 pages) (2009)
75. R.J.E. Merry, M.J.C. Ronde, R. van de Molengraft, K.R. Koops, M. Steinbuch, Directional repetitive control of a metrological AFM. IEEE Trans. Control Sys. Technol. **19**(6), 1622–1629 (2011)
76. B.J. Kenton, K.K. Leang, Flexure design using metal matrix composite materials: nanopositioning example, in *IEEE International Conference on Robotics and Automation (ICRA)* (2012)
77. K.K. Chew, M. Tomizuka, Digital control of repetitive errors in disk drive systems. IEEE Control Syst. Mag. **10**(1), 16–20 (1990)
78. M. Steinbuch, S. Weiland, T. Singh, Design of noise and period-time robust high-order repetitive control, with application to optical storage. Automatica **43**(12), 2086–2095 (2007)
79. Y. Shan, K.K. Leang, Dual-stage repetitive control with Prandtl-Ishlinskii hysteresis inversion for piezo-based nanopositioning. Mechatronics **22**, 271–281 (2012)
80. Y. Shan, K.K. Leang, Accounting for hysteresis in repetitive control design: nanopositioning example. Automatica **48**(8), 1751–1758 (2012)
81. K.K. Leang, Q. Zou, S. Devasia, Feedforward control of piezoactuators in atomic force microscope systems: inversion-based compensation for dynamics and hysteresis. IEEE Control Syst. Mag. **29**(1), 70–82 (2009, Special issue on Hysteresis)
82. E. Bayo, A finite-element approach to control the end-point motion of a single-link flexible robot. J. Robot. Syst. **4**, 63–75 (1987)
83. S. Devasia, D. Chen, B. Paden, Nonlinear inversion-based output tracking. IEEE Trans. Autom. Control **41**(7), 930–942 (1996)
84. Q. Zou, S. Devasia, Preview-based stable-inversion for output tracking. ASME J. Dyn. Syst. Meas. Control **121**(4), 625–630 (1999)
85. J.S. Dewey, K.K. Leang, S. Devasia, Experimental and theoretical results in output-trajectory redesign for flexible structures. ASME J. Dyn. Syst., Meas. Control **120**(4), 456–461 (1998)
86. Q. Zou, S. Devasia, Preview-based optimal inversion for output tracking: application to scanning tunneling microscopy. IEEE Control Sys. Technol. **12**(3), 375–386 (2004)
87. A.J. Fleming, A.G. Wills Optimal periodic trajectories for band-limited systems. IEEE Control Sys. Technol. **17**(3), 552–562 (2009)
88. G. Schitter, P.J. Thurner, P.K. Hansma, Design and input-shaping control of a novel scanner for high-speed atomic force microscopy. Mechatronics **18**(5–6), 282–288 (2008)
89. R.J.E. Merry, R. van de Molengraft, M. Steinbuch, Modeling of a walking piezo actuator. Sensors Actuators A Phys. **162**, 51–60 (2010)
90. Y. Zhao, S. Jayasuriya, Feedforward controllers and tracking accuracy in the presence of plant uncertainty. ASME J. Dyn. Syst. Meas. Control **117**, 490–495 (1995)
91. G. Schitter, A. Stemmer, F. Allgöwer, Robust 2DOF-control of a piezoelectric tube scanner for high-speed atomic force microscopy, in *American Control Conference* (2003), pp. 3720–3725
92. Y. Li, J. Bechhoefer, Feedforward control of closed-loop piezoelectric translation stage for atomic force microscope. Rev. Sci. Instrum. **78**(1), 013702 (2007)
93. S.S. Aphale, S. Devasia, S.O.R. Moheimani, High-bandwidth control of a piezoelectric nanopositioning stage in the presence of plant uncertainties. Nanotechnology **19**(12), 125503 (2008)
94. Y. Wu, Q. Zou, Robust inversion-based 2-DOF control design for output tracking: piezoelectric-actuator example. IEEE Control Sys. Technol. **17**(5), 1069–1082 (2009)
95. Y. Wu, Q. Zou, Iterative control approach to compensate for both the hysteresis and the dynamics effects of piezo actuators. IEEE Control Sys. Technol. **15**(5), 936–944 (2007)
96. D.A. Bristow, J. Dong, A.G. Alleyne, P. Ferriera, S. Salapaka, High bandwidth control of precision motion instrumentation. Rev. Sci. Instrum. **79**, 103704 (2008)

97. J. Ghosh, B. Paden, A pseudoinverse-based iterative learning control. IEEE Trans. Autom. Control **47**(5), 831–837 (2002)
98. Y. Li, J. Bechhoefer, Feedforward control of a piezoelectric flexure stage for AFM, in *American Control Conference* (2008), pp. 2703–2709
99. G.M. Clayton, S. Devasia, Image-based compensation of dynamic effects in scanning tunneling microscopes. Nanotechnology **16**, 809–818 (2005)
100. K.S. Kim, Q. Zou, Model-less inversion-based iterative control for output tracking: piezo actuator example. in *American Control Conference* (2008), pp. 2710–2715
101. M. Brokate, J. Sprekels, *Hysteresis and Phase Transitions* (Springer, New York, 1996)
102. A. Pantazi, A. Sebastian, H. Pozidis, E. Eleftheriou, Two-sensor-based $\mathscr{H}_\infty$ control for nanopositioning in probe storage, in *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference* (2005), pp. 1174–1179
103. J. Dong, S.M. Salapaka, P.M. Ferreira, Robust mimo control of a parallel kinematics nano-positioner for high resolution high bandwidth tracking and repetitive tasks, in *IEEE Conference on Decision and Control* (2007), pp. 4495–4500
104. Y. Wu, J. Shi, C. Su, Q. Zou, A control approach to cross-coupling compensation of piezotube scanners in tapping-mode atomic force microscope imaging. Rev. Sci. Instrum. **80**(4), 0433709 (2009)
105. S. Bashash, N. Jalili, Robust adaptive control of coupled parallel piezo-flexural nanopositioning stages. IEEE/ASME Trans. Mechatron. **14**(1), 11–20 (2009)
106. R.H. Comstock, Charge control of piezoelectric actuators to reduce hysteresis effects. US Patent 4,263,527, 1979
107. H. Kaizuka, B. Siu, A simple way to reduce hysteresis and creep when using piezoelectric actuators. Jpn. J. Appl. Phys. **27**(5), 773–776 (1988)
108. J.A. Main, E. Garcia, D.V. Netwon, Precision position control of piezoelectric actuators using charge feedback. AIAA J. Guid. Control. Dyn. **18**(5), 1068–1073 (1995)
109. K. Furutani, M. Urushibata, N. Mohri, Displacement control of piezoelectric element by feedback of induced charge. Nanotechnology **9**, 93–98 (1998)
110. A.J. Fleming, S.O.R. Moheimani, Precision current and charge amplifiers for driving highly capacitive piezoelectric loads. IEE Electron. Lett. **39**(3), 282–284 (2003)
111. A.J. Fleming, S.O.R. Moheimani, A grounded-load charge amplifier for reducing hysteresis in piezoelectric tube scanners. Rev. Sci. Instrum. **76**, 073707–1 (2005)
112. A.J. Fleming, K.K. Leang, Evaluation of charge drives for scanning probe microscope positioning stages, in *American Control Conference, Invited session on Advanced Mechanism Design, Modeling, and Control of SPMs* (2008), pp. 2028–2033
113. G.M. Clayton, T. Tien, A.J. Fleming, S.O.R. Moheimani, S. Devasia, Inverse-feedforward of charge controlled piezopositioners. Mechatronics **18**, 273–281 (2008)
114. B. Mokaberi, A.A.G. Requicha, Drift compensation for automatic nanomanipulation with scanning probe microscopes. IEEE Trans. Autom. Sci. Eng. **3**(3), 199–207 (2004)
115. Q. Yang, S. Jaqannathan, E.W. Bohannan, Atomic force microscope-based nanomanipulation with drift compensation. Int. J. Nanotechnol. **3**(4), 527–544 (2006)
116. V.Y. Yurov, A.N. Klimov, Scanning tunneling microscope calibration and reconstruction of real image: drift and slope elimination. Rev. Sci. Instrum. **65**(5), 1551–1557 (1994)
117. R. Staub, D. Alliata, C.Nicolini, Drift elimination in the calibration of scanning probe microscopes. Rev. Sci. Instrum. **66**(3), 2513–2516 (1995)
118. J.T. Woodward, D.K. Schwartz, Removing drift from scanning probe microscope images of periodic samples. J. Vac. Sci. Technol. B **16**(1), 51–53 (1998)
119. S.H. Huerth, H.D. Hallen, Quantitative method of image analysis when drift is present in a scanning probe microscope. J. Vac. Sci. Technol. B **21**(2), 714–718 (2000)

Chapter 8
Position Sensors for Nanopositioning

Andrew J. Fleming and Kam K. Leang

Abstract Position sensors with nanometer resolution are a key component of many precision imaging and fabrication machines. Since the sensor characteristics can define the linearity, resolution, and speed of the machine, the sensor performance is a foremost consideration. The first goal of this article is to define concise performance metrics and to provide exact and approximate expressions for error sources including nonlinearity, drift, and noise. The second goal is to review current position sensor technologies and to compare their performance. The sensors considered include: resistive, piezoelectric, and piezoresistive strain sensors; capacitive sensors; electrothermal sensors; eddy-current sensors; linear variable displacement transformers; interferometers; and linear encoders.

8.1 Introduction

The sensor requirements of a nanopositioning system are among the most demanding of any control system. The sensors must be compact, high-speed, immune to environmental variation, and able to resolve position down to the atomic scale. In many applications, such as atomic force microscopy [1, 2] or nanofabrication [3, 4], the performance of the machine or process is primarily dependent on the performance of the position sensor, thus, sensor optimization is a foremost consideration.

In order to define the performance of a position sensor, it is necessary to have strict definitions for the characteristics of interest. At present, terms such as accuracy, precision, nonlinearity, and resolution are defined loosely and often vary between manufacturers and researchers. The lack of a universal standard makes it

A.J. Fleming (✉)
University of Newcastle, Callaghan, NSW, Australia
e-mail: andrew.fleming@newcastle.edu.au

K.K. Leang
University of Utah, Salt Lake City, UT, USA
e-mail: kam.k.leang@utah.edu

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_8

difficult to predict the performance of a particular sensor from a set of specifications. Furthermore, specifications may not be in a form that permits the prediction of closed-loop performance.

This article provides concise definitions for the linearity, drift, bandwidth, and resolution of position sensors. The measurement errors resulting from each source are then quantified and bounded to permit a straightforward comparison between sensors. An emphasis is placed on specifications that allow the prediction of closed-loop performance as a function of the controller bandwidth.

Although there are presently no international standards for the measurement or reporting of position sensor performance, this article is aligned with the definitions and methods reported in the ISO/IEC 98:1993 Guide to the Expression of Uncertainty in Measurement [5], and the ISO 5723 Standard on Accuracy (Trueness and Precision) of Measurement Methods and Results [6].

The noise and resolution of a position sensor is potentially one of the most misreported sensor characteristics. The resolution is commonly reported without mention of the bandwidth or statistical definition and thus has little practical value. To improve the understanding of this issue, the relevant theory of stochastic processes is reviewed in Sect. 8.2. The variance is then utilized to define a concise statistical description of the resolution, which is a straightforward function of the noise density, bandwidth, and $1/f$ corner frequency.

The second goal of this article is to provide a tutorial introduction and comparison of sensor technologies suitable for nanopositioning applications. To be eligible for inclusion, a sensor must be capable of a 6σ-resolution better than 10 nm with a bandwidth greater than 10 Hz. The sensor can not introduce friction or contact forces between the reference and moving target, or exhibit hysteresis or other characteristics that limit repeatability.

The simplest sensor considered is the metal foil strain gauge discussed in Sect. 8.3.1. These devices are often used for closed-loop control of piezoelectric actuators but are limited by temperature dependence and low sensitivity [7]. Piezoresistive and piezoelectric strain sensors provide improved sensitivity but at the cost of stability and DC performance.

The most commonly used sensors in nanopositioning systems [8] are the capacitive and eddy-current sensors discussed in Sects. 8.3.4 and 8.3.6. Capacitive and eddy-current sensors are more complex than strain sensors but can be designed with sub-nanometer resolution, albeit with comparably small range and low bandwidth. They are used extensively in applications such as atomic force microscopy [2, 9–11] and nanofabrication [4, 12]. The linear variable displacement transformer (LVDT) described in Sect. 8.3.7 is a similar technology that is intrinsically linear. However, this type of sensor is larger than a capacitive sensor and due to the larger range, is not as sensitive.

To achieve high absolute accuracy over a large range, the reference standard is the laser heterodyne interferometer discussed in Sect. 8.3.8. Although bulky and costly, the interferometer has been the sensor of choice for applications such as IC wafer steppers [13, 14] and metrological systems [15]. New fiber interferometers are also discussed that are extremely compact and ideal for extreme environments.

Aside from the cost and size, the foremost difficulties associated with an interferometer are the susceptibility to beam interference, variation in the optical medium, and alignment error. Since an interferometer is an incremental position sensor, if the beam is broken or the maximum traversing speed is exceeded, the system must be returned to a known reference before continuing. These difficulties are somewhat alleviated by the absolute position encoders described in Sect. 8.3.9. A position encoder has a read-head that is sensitive to a geometric pattern encoded on a reference scale. Reference scales operating on the principle of optical interference can have periods of 128 nm and a resolution of a few nanometers.

Other sensor technologies that were considered but did not fully satisfy the eligibility criteria include optical triangulation sensors [16], Hall effect sensors, and magnetoresistive sensors. In general, optical triangulation sensors are available in ranges from 0.5 mm to 1 m with a maximum resolution of approximately 100 nm. Hall effect sensors are sensitive to magnetic field strength and hence the distance from a known magnetic source. These sensors have a high resolution, large range, and wide bandwidth but are sensitive to external magnetic fields and exhibit hysteresis of up to 0.5 % which degrades the repeatability. The magnetoresistive sensor is similar except that the resistance, rather than the induced voltage, is sensitive to magnetic field. Although typical anisotropic magnetoresistive (AMR) sensors offer similar characteristics to the Hall effect sensor, recent advances stimulated by the hard disk industry have provided major improvements [17]. In particular, the giant magnetoresistive effect (GMR) can exhibit two orders of magnitude greater sensitivity than the AMR effect which equates to a resistance change of up to 70 % at saturation. Such devices can also be miniaturized and are compatible with lithographic processes. Packaged GMR sensors in a full-bridge configuration are now available from NVE Corporation, NXP Semiconductor, Siemens, and Sony. Aside from the inherent nonlinearities associated with the magnetic field, the major remaining drawback is the hysteresis of up to 4 % which can severely impact the performance in nanopositioning applications. Despite this, miniature GMR sensors have shown promise in nanopositioning applications by keeping the changes in magnetic field small [18, 19]. However, to date, the linearity and hysteresis of this approach has not been reported.

8.2 Sensor Characteristics

8.2.1 Calibration and Nonlinearity

Position sensors are designed to produce an output that is directly proportional to the measured position. However, in reality, all position sensors have an unknown offset, sensitivity, and nonlinearity. These effects must be measured and accounted for in order to minimize the uncertainty in position.

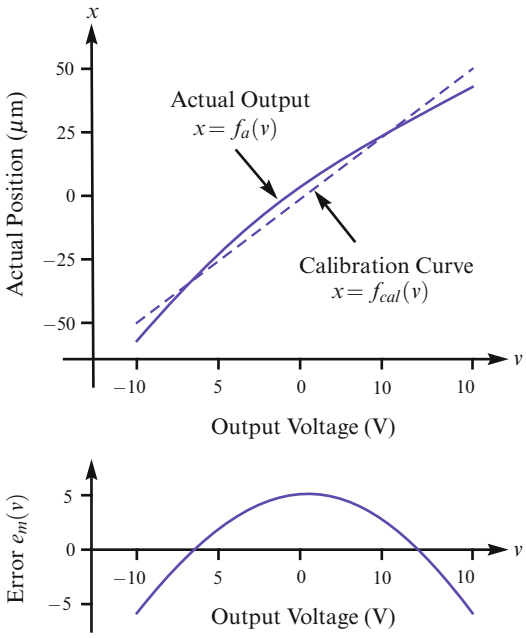


Fig. 8.1 The actual position versus the output voltage of a position sensor. The calibration function $f_{\text{cal}}(v)$ is an approximation of the sensor mapping function $f_a(v)$ where v is the voltage resulting from a displacement x. $e_m(v)$ is the residual error

The typical output voltage curve for a capacitive position sensor is illustrated in Fig. 8.1. A nonlinear function $f_a(v)$ maps the output voltage v to the actual position x. The calibration process involves finding a curve $f_{\text{cal}}(v)$ that minimizes the mean-square error, known as the least-squares fit, defined by

$$\theta^* = \arg\min \sum_{i=1}^{N} [x_i - f_{\text{cal}}(\theta, v_i)]^2 , \tag{8.1}$$

where v_i and x_i are the data points and θ^* is the vector of optimal parameters for $f_{\text{cal}}(\theta, v)$. The simplest calibration curve, as shown in Fig. 8.1, is a straight line of best fit,

$$f_{\text{cal}}(v) = \theta_0 + \theta_1 v. \tag{8.2}$$

In the above equation, the sensor offset is θ_0 and the sensitivity is θ_1 μm/V. More complex mapping functions are also commonly used, including the higher order polynomials

$$f_{\text{cal}}(v) = \theta_0 + \theta_1 v + \theta_2 v^2 + \theta_3 v^3 \cdots . \tag{8.3}$$

Once the calibration function $f_{\text{cal}}(v)$ is determined, the actual position can be estimated from the measured sensor voltage. Since the calibration function does not perfectly describe the actual mapping function $f_a(v)$, a mapping error results. The mapping error $e_m(v)$ is the residual of (8.1), that is

$$e_m(v) = f_a(v) - f_{\text{cal}}(\theta^*, v). \tag{8.4}$$

If $e_m(v)$ is positive, the true position is greater than the estimated value and vice versa. Although the mapping error has previously been defined as the peak-to-peak variation of $e_m(v)$ [20], this may underestimate the positioning error if $e_m(v)$ is not symmetric. A more conservative definition of the mapping error (e_m) is

$$e_m = \pm \max |e_m(v)| . \tag{8.5}$$

It is also possible to specify an unsymmetrical mapping error such as $+\max e_m(v)$, $-\min e_m(v)$ however this is more complicated. For the sake of comparison, the maximum mapping error (nonlinearity) is often quoted as a percentage of the full-scale range (FSR), for example

$$\text{Mapping Error}(\%) = \pm 100 \frac{\max |e_m(v)|}{\text{FSR}}. \tag{8.6}$$

Since there is no exact consensus on the reporting of nonlinearity, it is important to know how the mapping error is defined when evaluating the specifications of a position sensor. A less conservative definition than that stated above may exaggerate the accuracy of a sensor and lead to unexplainable position errors. It may also be necessary to consider other types of nonlinearity such as hysteresis [21]. However, sensors that exhibit hysteresis have poor repeatability and are generally not considered for precision sensing applications.

8.2.2 Drift and Stability

In addition to the nonlinearity error discussed above, the accuracy of a positioning sensor can also be severely affected by changes in the mapping function $f_a(v)$. The parameters of $f_a(v)$ may drift over time, or be dependent on environmental conditions such as temperature, humidity, dust, or gas composition. Although, the actual parametric changes in $f_a(v)$ can be complicated, it is possible to bound the variations by an uncertainty in the sensitivity and offset. That is,

$$f_a(v) = (1 + k_s) f_a^*(v) + k_o, \tag{8.7}$$

where k_s is the sensitivity variation usually expressed as a percentage, k_o is the offset variation, and $f_a^*(v)$ is the nominal mapping function at the time of calibration. With the inclusion of sensitivity variation and offset drift, the mapping error is

$$e_d(v) = (1 + k_s) f_a^*(v) + k_o - f_{\text{cal}}(v). \tag{8.8}$$

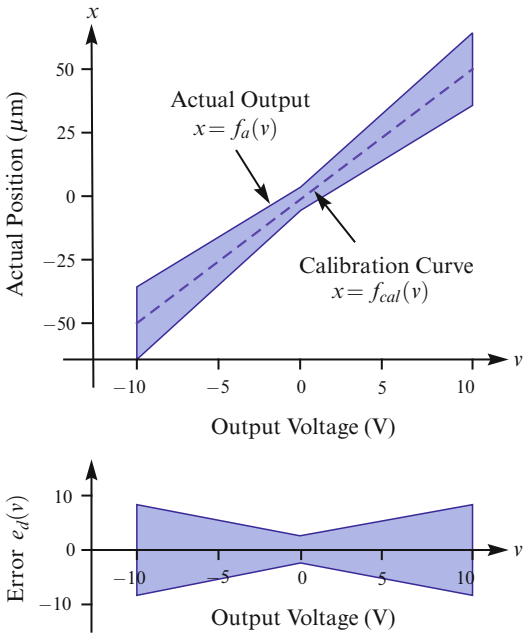


Fig. 8.2 The worst-case range of a linear mapping function $f_a(v)$ for a given error in sensitivity and offset. In this example the greatest error occurs at the maximum and minimum of the range

Equations (8.7) and (8.8) are illustrated graphically in Fig. 8.2. If the nominal mapping error is assumed to be small, the expression for error can be simplified to

$$e_d(v) = k_s f_{\text{cal}}(v) + k_o. \tag{8.9}$$

That is, the maximum error due to drift is

$$e_d = \pm \left(k_s \max |f_{\text{cal}}(v)| + k_o\right). \tag{8.10}$$

Alternatively, if the nominal calibration can not be neglected or if the shape of the mapping function actually varies with time, the maximum error due to drift must be evaluated by finding the worst-case mapping error defined in (8.5).

8.2.3 Bandwidth

The bandwidth of a position sensor is the frequency at which the magnitude of the transfer function $v(s)/x(s)$ drops by 3 dB. Although the bandwidth specification is useful for predicting the resolution of a sensor, it reveals very little about the measurement errors caused by sensor dynamics. For example, a sensor phase-lag of only 12 degrees causes a measurement error of 10 % FSR.

If the sensitivity and offset have been accounted for, the frequency domain position error is

$$e_{\mathrm{bw}}(s) = x(s) - v(s), \tag{8.11}$$

which is equal to

$$e_{\mathrm{bw}}(s) = x(s)\left(1 - P(s)\right), \tag{8.12}$$

where $P(s)$ is the sensor transfer function and $(1 - P(s))$ is the multiplicative error. If the actual position is a sine wave of peak amplitude A, the maximum error is

$$e_{\mathrm{bw}} = \pm A\left|1 - P(s)\right|. \tag{8.13}$$

The worst-case error occurs when $A = \mathrm{FSR}/2$, in this case,

$$e_{\mathrm{bw}} = \pm\frac{\mathrm{FSR}}{2}\left|1 - P(s)\right|. \tag{8.14}$$

The error resulting from a Butterworth response is plotted against normalized frequency in Fig. 8.3. Counter to intuition, the higher order filters produce more error, which is surprising because these filters have faster roll-off, however, they also contribute more phase-lag. If the poles of the filter are assumed to be equal to the cut-off frequency, the low-frequency magnitude of $|1 - P(s)|$ is approximately

$$|1 - P(s)| \approx n\frac{f}{f_c}, \tag{8.15}$$

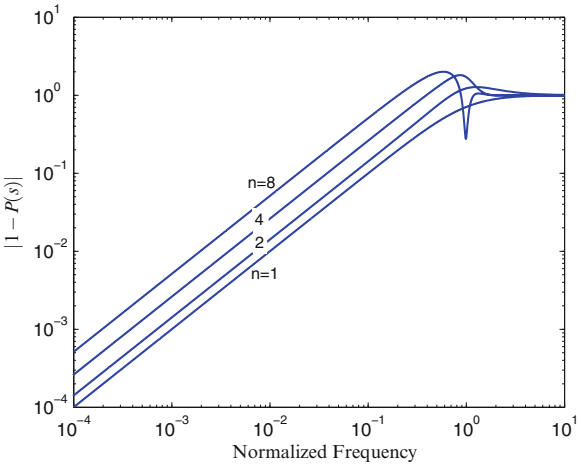


Fig. 8.3 The magnitude of error caused by the sensor dynamics $P(s)$. The frequency axis is normalized to the sensor 3 dB bandwidth. Lower order sensor dynamics result in lower error but typically result in significantly lesser bandwidths. In this example the dynamics are assumed to be nth order Butterworth

where n is the filter order and f_c is the bandwidth. The resulting error is approximately

$$e_{\text{bw}} \approx \pm A\, n \frac{f}{f_c}. \tag{8.16}$$

That is, the error is proportional to the magnitude of the signal, filter order, and normalized frequency. This is significant because the sensor bandwidth must be significantly higher than the operating frequency if dynamic errors are to be avoided. For example, if an absolute accuracy of 10 nm is required when measuring a signal with an amplitude of 100 μm, the sensor bandwidth must be ten-thousand times greater than the signal frequency.

In the above derivation, the position signal was assumed to be sinusoidal, for different trajectories, the maximum error must be found by simulating Eq. (8.12). Although the RMS error can be found analytically by applying Parseval's equality, there is no straightforward method for determining the peak error, aside from numerical simulation. In general, signals that contain high-frequency components such as square and triangle waves cause the greatest peak error.

8.2.4 Noise

In addition to the actual position signal, all sensors produce some additive measurement noise. In many types of sensor, the majority of noise arises from the thermal noise in resistors and the voltage and current noise in conditioning circuit transistors. As these noise processes can be approximated by Gaussian random processes, the total measurement noise can also be approximated by a Gaussian random process.

A Gaussian random process produces a signal with normally distributed values that are correlated between instances of time. We also assume that the noise process is zero-mean and that the statistical properties do not change with time, that is, the noise process is stationary. A Gaussian noise process can be described by either the autocorrelation function or the power spectral density. The autocorrelation function of a random process $\mathscr{X}$ is

$$R_{\mathscr{X}}(\tau) = E\left[\mathscr{X}(t)\mathscr{X}(t+\tau)\right], \tag{8.17}$$

where E is the expected value operator. The autocorrelation function describes the correlation between two samples separated in time by τ. Of special interest is $R_{\mathscr{X}}(0)$ which is the variance of the process. The variance of a signal is the expected value of the varying part squared. That is,

$$\text{Var}\ \mathscr{X} = E\left[(\mathscr{X} - E[\mathscr{X}])^2\right]. \tag{8.18}$$

Another term used to quantify the dispersion of a random process is the standard deviation σ which is the square-root of variance,

$$\sigma_{\mathscr{X}} = \text{standard deviation of } \mathscr{X} = \sqrt{\text{Var } \mathscr{X}}. \tag{8.19}$$

The standard deviation is also the root-mean-square (RMS) value of a zero-mean random process.

The power spectral density $S_{\mathscr{X}}(f)$ of a random process represents the distribution of power or variance across frequency f. For example, if the random process under consideration was measured in Volts, the power spectral density would have the units of V^2/Hz. The power spectral density can be found by either the averaged periodogram technique, or from the autocorrelation function. The periodogram technique involves averaging a large number of Fourier transforms of a random process,

$$2 \times E\left[\frac{1}{T}\left|\mathscr{F}\left\{\mathscr{X}_T(t)\right\}\right|^2\right] \Rightarrow S_{\mathscr{X}}(f) \text{ as} T \Rightarrow \infty. \tag{8.20}$$

This approximation becomes more accurate as T becomes larger and more records are used to compute the expectation. In practice, $S_{\mathscr{X}}(f)$ is best measured using a spectrum or network analyzer, these devices compute the approximation progressively so that large time records are not required. The power spectral density can also be computed from the autocorrelation function. The relationship between the autocorrelation function and power spectral density is known as the Wiener–Khinchin relations, given by

$$S_{\mathscr{X}}(f) = 2\mathscr{F}\left\{R_{\mathscr{X}}(\tau)\right\} = 2\int_{-\infty}^{\infty} R_{\mathscr{X}}(\tau)e^{-j2\pi f\tau}\, d\tau, \text{ and} \tag{8.21}$$

$$R_{\mathscr{X}}(\tau) = \frac{1}{2}\mathscr{F}^{-1}\left\{S_{\mathscr{X}}(f)\right\} = \frac{1}{2}\int_{-\infty}^{\infty} S_{\mathscr{X}}(f)e^{j2\pi f\tau}\, df. \tag{8.22}$$

If the power spectral density is known, the variance of the generating process can be found from the area under the curve, that is

$$\sigma_{\mathscr{X}}^2 = E\left[\mathscr{X}^2(t)\right] = R_{\mathscr{X}}(0) = \int_0^{\infty} S_{\mathscr{X}}(f)\, df. \tag{8.23}$$

Rather than plotting the frequency distribution of power or variance, it is often convenient to plot the frequency distribution of the standard deviation, which is referred to as the spectral density. It is related to the standard power spectral density function by a square-root, that is,

$$\text{Spectral Density} = \sqrt{S_{\mathscr{X}}(f)}. \tag{8.24}$$

The units of $\sqrt{S_{\mathscr{X}}(f)}$ are units$/\sqrt{\text{Hz}}$ rather than units^2/Hz. The spectral density is preferred in the electronics literature as the RMS value of a noise process can be determined directly from the noise density and effective bandwidth. For example, if the noise density is a constant c $V/\sqrt{\text{Hz}}$ and the process is perfectly band limited to f_c Hz, the RMS value or standard deviation of the resulting signal is $c\sqrt{f_c}$. To distinguish between power spectral density and noise density, A is used for power spectral density and $\sqrt{A}$ is used for noise density. An advantage of the spectral density is that a gain k applied to a signal $u(t)$ also scales the spectral density by k. This differs from the standard power spectral density function that must be scaled by k^2.

Since the noise in position sensors is primarily due to thermal noise and $1/f$ (flicker) noise, the power spectral density can be approximated by

$$S(f) = A\frac{f_{\text{nc}}}{|f|} + A, \tag{8.25}$$

where A is the power spectral density and f_{nc} is the noise corner frequency illustrated in Fig. 8.4. The variance of this process can be found by evaluating Eq. (8.23). That is,

$$\sigma^2 = \int_{f_l}^{f_h} A\frac{f_{\text{nc}}}{|f|} + A\, df. \tag{8.26}$$

where f_l and f_h define the bandwidth of interest. Extremely low-frequency noise components are considered to be drift. In positioning applications, f_l is typically chosen between 0.01 and 0.1 Hz. By solving Eq. (8.26), the variance is

$$\sigma^2 = Af_{\text{nc}} \ln\frac{f_h}{f_l} + A(f_h - f_l). \tag{8.27}$$

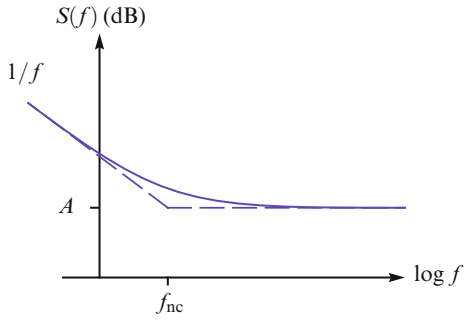


Fig. 8.4 A constant power spectral density that exhibits $1/f$ noise at low frequencies. The *dashed lines* indicate the asymptotes

If the upper-frequency limit is due to a linear filter and $f_h >> f_l$, the variance can be modified to account for the finite roll-off of the filter, that is

$$\sigma^2 = A f_{\text{nc}} \ln \frac{f_h}{f_l} + A k_e f_h. \tag{8.28}$$

where k_e is a correction factor that accounts for the finite roll-off. For a first, second, third, and fourth order response k_e is equal to 1.57, 1.11, 1.05, and 1.03, respectively [22].

8.2.5 Resolution

The random noise of a position sensor causes an uncertainty in the measured position. If the distance between two measured locations is smaller than the uncertainty, it is possible to mistake one point for the other. In fabrication and imaging applications, this can cause manufacturing faults or imaging artifacts. To avoid these eventualities, it is critical to know the minimum distance between two adjacent but unique locations.

Since the random noise of a position sensor has a potentially large dispersion, it is impractically conservative to specify a resolution where adjacent locations never overlap. Instead, it is preferable to state the probability that the measured value is within a certain error bound. Consider the plot of three noisy measurements in Fig. 8.5 where the resolution δ_y is shaded in gray. The majority of sample points

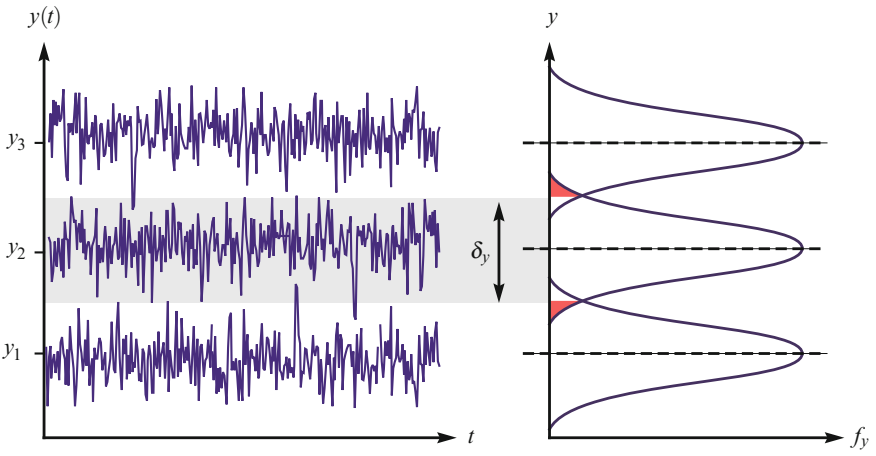


Fig. 8.5 The time-domain recording $y(t)$ of a position sensor at three discrete positions y_1, y_2, and y_3. The large *shaded area* represents the resolution of the sensor and the approximate peak-to-peak noise of the sensor. The probability density function f_y of each signal is shown on the *right*

in y_2 fall within the bound $y_2 \pm \delta_y/2$. However, not all of the samples of y_2 lie within the resolution bound, as illustrated by the overlap of the probability density functions. To find the maximum measurement error, the resolution is added to other error sources as described in Sect. 8.2.6.

If the measurement noise is approximately Gaussian distributed, the resolution can be quantified by the standard deviation σ (RMS value) of the noise. The empirical rule [23] states that there is a 99.7 % probability that a sample of a Gaussian random process lie within $\pm 3\sigma$. Thus, if we define the resolution as $\delta = 6\sigma$ there is only a 0.3 % probability that a sample lies outside of the specified range. To be precise, this definition of resolution is referred to as the 6σ-resolution. Beneficially, no statistical measurements are required to obtain the 6σ-resolution if the noise is Gaussian distributed.

In other applications where more or less overlap between points is tolerable, another definition of resolution may be more appropriate. For example, the 4σ-resolution would result in an overlap 4.5 % of the time, while the 10σ-resolution would almost eliminate the probability of an overlap. Thus, it is not the exact definition that is important; rather, it is the necessity of quoting the resolution together with its statistical definition.

Although there is no international standard for the measurement or reporting of resolution in a positioning system, the ISO 5725 Standard on Accuracy (Trueness and Precision) of Measurement Methods and Results [6] defines precision as the standard deviation (RMS Value) of a measurement. Thus, the 6σ-resolution is equivalent to six times the ISO definition for precision.

If the noise is not Gaussian distributed, the resolution can be measured by obtaining the 99.7 percentile bound directly from a time-domain recording. To obtain a statistically valid estimate of the resolution, the recommended recording length is 100 s with a sampling rate 15 times the sensor bandwidth [24]. An anti-aliasing filter is required with a cut-off frequency 7.5 times the bandwidth. Since the signal is likely to have a small amplitude and large offset, an AC coupled preamplifier is required with a high-pass cut-off of 0.03 Hz or lower [25].

Another important parameter that must be specified when quoting resolution is the sensor bandwidth. In Eq. (8.28), the variance of a noise process is shown to be approximately proportional to the bandwidth f_h. By combining Eq. (8.28) with the above definition of resolution, the 6σ-resolution can be found as a function of the bandwidth f_h, noise density $\sqrt{A}$, and $1/f$ corner frequency f_{nc},

$$6\sigma\text{−resolution} = 6\sqrt{A}\sqrt{f_{\text{nc}} \ln \frac{f_h}{f_l} + k_e f_h}. \tag{8.29}$$

From Eq. (8.29), it can be observed that the resolution is approximately proportional to the square-root of bandwidth when $f_h >> f_{nc}$. It is also clear that the $1/f$ corner frequency limits the improvement that can be achieved by reducing the bandwidth. Note that Eq. (8.29) relies on a noise spectrum of the form (8.25) which may not adequately represent some sensors. The resolution of sensors with irregular

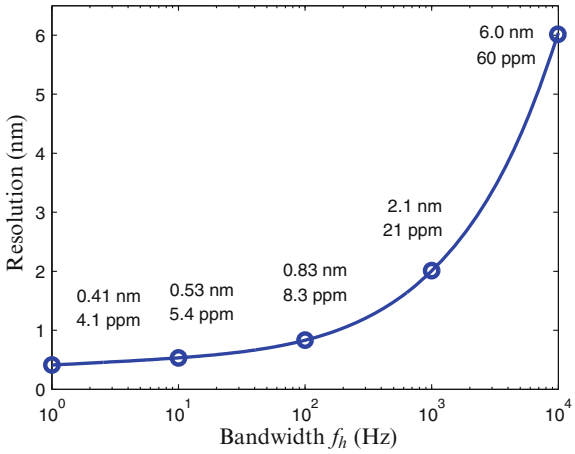


Fig. 8.6 The resolution versus bandwidth of a position sensor with a noise density of 10 pm/$\sqrt{\text{Hz}}$ and a $1/f$ corner frequency of 10 Hz ($f_l = 0.01$ Hz and $k_e = 1$). At low frequencies, the noise is dominated by $1/f$ noise; however, at high frequencies, the noise increases by a factor of 3.16 for every decade of bandwidth

spectrum's can be found by solving (8.23) numerically. Alternatively, the resolution can evaluated from time-domain data, as discussed above.

The trade-off between resolution and bandwidth can be illustrated by considering a typical position sensor with a range of 100 μm, a noise density of 10 pm/$\sqrt{\text{Hz}}$, and a $1/f$ corner frequency of 10 Hz. The resolution is plotted against bandwidth in Fig. 8.6. When the bandwidth is below 100 Hz, the resolution is dominated by $1/f$ noise. For example, the resolution is only improved by a factor of two when the bandwidth is reduced by a factor of 100. Above 1 kHz, the resolution is dominated by the flat part of the power spectral density, thus a ten times increase in bandwidth from 1 to 10 kHz causes an approximately $\sqrt{10}$ reduction in resolution.

Many types of position sensors have a limited full-scale range (FSR); examples include strain sensors, capacitive sensors, and inductive sensors. In this class of sensor, sensors of the same type and construction tend to have an approximately proportional relationship between the resolution and range. As a result, it is convenient to consider the ratio of resolution to the full-scale range, or equivalently, the dynamic range (DNR). This figure can be used to quickly estimate the resolution from a given range, or conversely, to determine the maximum range given a certain resolution. A convenient method for reporting this ratio is in parts-per-million (ppm), that is

$$\text{DNR}_{\text{ppm}} = 10^6 \frac{6\sigma\text{−resolution}}{\text{Full scale range}}. \tag{8.30}$$

This measure is equivalent to the resolution in nanometers of a sensor with a range of 1 mm. In Fig. 8.6 the resolution is reported in terms of both absolute distance and the dynamic range in ppm. The dynamic range can also be stated in decibels,

$$\text{DNR}_{\text{db}} = 20 \log_{10} \frac{\text{Full scale range}}{6\sigma\text{−resolution}}. \tag{8.31}$$

Table 8.1 Summary of the exact and simplified worst-case measurement errors

Error source	Exact	Simplified bound
Mapping error e_m	$f_a(v) - f_{\text{cal}}(\theta^*, v)$	$\pm \max \lvert e_m(v) \rvert$
Drift e_d	$(1 + k_s) f_a^*(v) + k_o - f_{\text{cal}}(v)$	$\pm (k_s \max \lvert f_{\text{cal}}(v) \rvert + k_o)$
Bandwidth e_{bw}	$\mathscr{F}^{-1}\{x(s)(1 - P(s))\}$	$\pm \frac{Anf}{f_c}$ (sine-wave)
Noise δ	NA	$6\sqrt{A}\sqrt{f_{\text{nc}} \ln \frac{f_h}{f_l} + k_e f_h}$

Due to the strong dependence of resolution and dynamic range on the bandwidth of interest, it is clear that these parameters cannot be reported without the frequency limits f_l and f_h, to do so would be meaningless. Even if the resolution is reported correctly, it is only relevant for a single operating condition. A better alternative is to report the noise density and $1/f$ corner frequency, which allows the resolution and dynamic range to be calculated for any operating condition. These parameters are also sufficient to predict the closed-loop noise of a positioning system that incorporates the sensor [26]. If the sensor noise is not approximately Gaussian or the spectrum is irregular, the resolution is measured using the process described above for a range of logarithmically spaced bandwidths.

8.2.6 Combining Errors

The exact and worst-case errors described in Sect. 8.2 are summarized in Table 8.1. In many circumstances, it is not practical to consider the exact error as this is dependent on the position. Rather, it is preferable to consider only the simplified worst-case error. An exception to the use of worst-case error is the drift-error e_d. In this case, it may be unnecessarily conservative to consider the maximum error since the exact error is easily related to the sensor output by the uncertainty in sensitivity and offset.

To calculate the worst-case error e_t, the individual worst-case errors are summed, that is

$$e_t = e_m + e_d + e_{\text{bw}} + \delta/2, \tag{8.32}$$

where e_m, e_d, e_{bw}, and $\delta/2$ are the mapping error, the drift error, the error due to finite bandwidth, and the error due to noise whose maximum is half the resolution δ. The sum of the mapping and drift error can be referred to as the static trueness error e_s which is the maximum error in a static position measurement when the noise is effectively eliminated by a slow averaging filter. The total error and the static trueness error are illustrated graphically in Fig. 8.7.

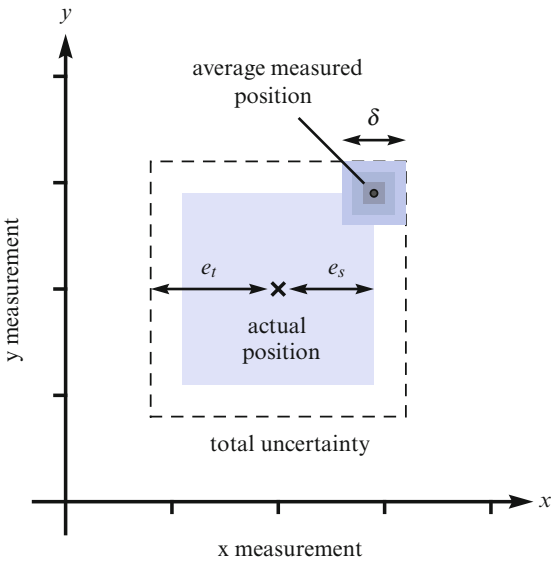


Fig. 8.7 The total uncertainty of a two-dimensional position measurement is illustrated by the *dashed box*. The total uncertainty e_t is due to both the static trueness error e_s and the noise δ

8.2.7 Metrological Traceability

The error of a position sensor has been evaluated with respect to the true position. However, in practice, the "true" position is obtained from a reference sensor that may also be subject to calibration errors, nonlinearity, and drift. If the tolerance of the calibration instrument is significant, this error must be included when evaluating the position sensor accuracy. However, such consideration is usually unnecessary as the tolerance of the calibration instrument is typically negligible compared to the position sensor being calibrated. To quantify the tolerance of a calibration instrument, it must be compared to a metrological reference for distance. Once the tolerance is known, measurements produced by the instrument can then be related directly to the reference, such measurements are said to be metrologically traceable.

Metrological traceability is defined as "the property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty" [27]. The reference for a distance measurement is the meter standard, defined by the distance traveled by light in vacuum over 1/299 792 458 s. Laser interferometers are readily calibrated to this standard since the laser frequency can be compared to the time standard which is known to an even higher accuracy than the speed of light.

Metrological traceability has little meaning by itself and must be quoted with an associated uncertainty to be valid [27]. If a position sensor is calibrated by an instrument that is metrologically traceable, subsequent measurements made by the position sensor are also metrologically traceable to within the bounds of the uncertainty for a specified operating environment [5].

To obtain metrologically traceable measurements with the least uncertainty, an instrument should be linked to the reference standard through the least number of intervening instruments or measurements. All countries have a national organization that maintains reference standards for the calibration instruments. It should be noted that these organizations have individual policies for the reporting of traceability if their name is quoted. For example, to report that a measurement is NIST Traceable, the policy of the National Institute of Standards and Technology (USA), must be adhered to. Examples of measurement standards organizations include:

- National Measurement Institute (NIM), Australia
- Bureau International des Poids et Mesures (BIPM), France
- Physikalisch-Technische Bundesanstalt (PTB), Germany
- National Metrology Institute of Japan (NMIJ), Japan
- British Standards Institution (BMI), United Kingdom
- National Institute of Standards and Technology (NIST), USA.

8.3 Nanometer Position Sensors

8.3.1 Resistive Strain Sensors

Due to their simplicity and low-cost, resistive strain gauges are widely used for position control of piezoelectric actuators. Resistive strain gauges can be integrated into the actuator or bonded to the actuator surface. An example of a piezoelectric actuator and resistive strain gauge is pictured in Fig. 8.14a. Other application examples can be found in references [28–31].

Resistive strain gauges are constructed from a thin layer of conducting foil laminated between two insulating layers. With a zig-zag conductor pattern, strain gauges can be designed for high sensitivity in only one direction, for example, elongation. When a strain gauge is elongated, the resistance increases proportionally. The change in resistance per unit strain is known as the gauge factor GF defined by

$$\mathrm{GF} = \frac{\Delta R/R_G}{\epsilon}, \tag{8.33}$$

where ΔR is the change in resistance from the nominal value R_G for a strain ϵ. As the gauge factor is typically in the order of 1 or 2, the change is resistance is similar in magnitude to the percentage of strain. For a piezoelectric transducer with a maximum strain of approximately 0.1 %, the change in resistance is around 0.1 %. This small variation requires a bridge circuit for accurate measurement.

In Fig. 8.14a, a 10 mm Noliac SCMAP07 piezoelectric actuator is pictured with a strain gauge bonded to each of the two non-electrode sides. The strain gauges are Omega SGD-3/350-LY13 gauges, with a nominal resistance of 350 Ohms and package dimensions of 7×4 mm. The electrical wiring of the strain gauges

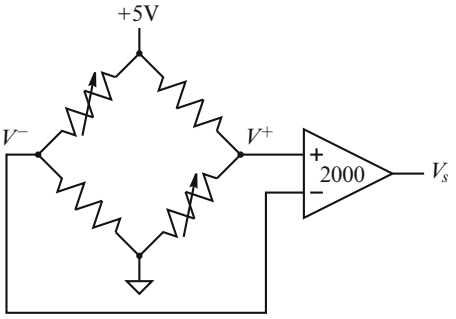


Fig. 8.8 A two-varying-element bridge circuit that contains two fixed resistors and two strain dependent resistors. All of the nominal resistance values are equal. A simultaneous change in the two-varying-elements produces a differential voltage across the bridge

is illustrated in Fig. 8.8. The two-varying-element bridge circuit is completed by two dummy 350 Ohm wire wound resistors and excited by a 5 Volt DC source. The differential bridge voltage ($V^+ - V^-$) is acquired and amplified by a Vishay Micro-Measurements 2120B strain gauge amplifier. The developed voltage from a two-varying-element bridge is

$$V_s = \frac{A_v V_b}{2}\left(\frac{\Delta R}{R_G + \Delta R/2}\right), \tag{8.34}$$

where $A_v = 2000$ is the differential gain and $V_b = 5$ V is the excitation voltage. By substituting (8.33) into (8.34) and neglecting the small bridge nonlinearity,[1] the measured voltage is proportional to the strain ϵ and displacement d by

$$V_s = \tfrac{1}{2} A_v V_b \mathrm{GF} \epsilon \tag{8.35}$$

$$V_s = \tfrac{1}{2L} A_v V_b \mathrm{GF} d, \tag{8.36}$$

where L is the actuator length. With a gauge factor of 1, the position sensitivity of the amplified strain sensor is predicted to be 0.5 V/μm which implies a full-scale voltage of 5 V from a displacement of 10 μm. The actual sensitivity was found to be 0.3633 V/μm [31].

The bridge configuration shown in Fig. 8.8 is known as the two-varying-element bridge. It has twice the sensitivity of a single-element bridge but is also slightly nonlinear and sensitive to temperature variations between the gauge and bridge resistances. A detailed review of bridge circuits and their associated instrumentation can be found in reference [32]. The best configuration is the four-varying-element differential bridge. This arrangement requires four strain gauges, two of which experience negative strain and another two that experience positive strain. Since the

[1] In a two-varying-element bridge circuit, the nonlinearity due to $\Delta R/2$ in Eq. (8.34) is 0.5 % nonlinearity per percent of strain [32]. Since the maximum strain of a piezoelectric actuator is 0.1 %, the maximum nonlinearity is only 0.05 % and can be neglected. If this magnitude of nonlinearity is not tolerable, compensating circuits are available [32].

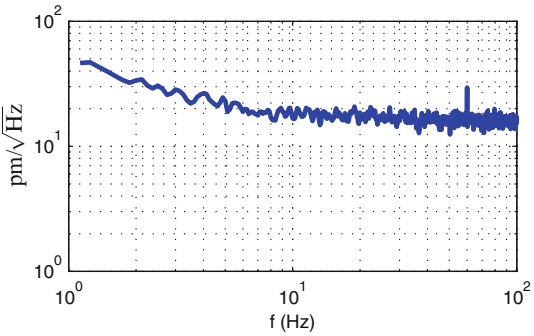


Fig. 8.9 The noise density of the strain sensor and instrumentation. The spectrum can be approximated by a constant spectral density and $1/f$ noise

bridge is made entirely from the same elements, the four-varying-element bridge is insensitive to temperature variation. The bridge nonlinearity is also eliminated. In applications where regions of positive and negative strain are not available, the two-varying-element bridge is used.

Compared to other position sensors, strain gauges are compact, low-cost, precise, and highly stable, particularly in a full-bridge configuration [30, 32]. However, a major disadvantage is the high measurement noise that arises from the resistive thermal noise and the low sensitivity. The power spectral density of the resistive thermal noise is

$$S(f) = 4kTR \ \ \text{V}^2/\text{Hz}, \tag{8.37}$$

where k is the Boltzmann constant (1.38×10^{-23}), T is the room temperature in Kelvin (300°), and R is the resistance of each element in the bridge. In addition to the thermal noise, the current through the bridge also causes $1/f$.

The strain gauge pictured in Fig. 8.14a has a resistance of 350 Ohms, hence the spectral density is $2.4\,\text{nV}/\sqrt{\text{Hz}}$. Since the sensitivity is $0.3633\,\text{V}/\mu\text{m}$, the predicted spectral density is $13\,\text{pm}/\sqrt{\text{Hz}}$. This figure agrees with the experimentally measured spectral density plotted in Fig. 8.9. The sensor exhibits a noise density of approximately $15\,\text{pm}/\sqrt{\text{Hz}}$ and a $1/f$ noise corner frequency of around 5 Hz. This compares poorly with the noise density of a typical inductive or capacitive sensor which is on the order of $1\,\text{pm}/\sqrt{\text{Hz}}$ for a range of $10\,\mu\text{m}$. Hence, strain gauges are rarely used in systems designed for high resolution. If they are utilized in such systems, the closed-loop bandwidth must be severely restrained.

As an example of strain gauge resolution, we consider a typical two-varying-element strain gauge with an excitation of 5 V and a gauge factor of 1. The full-scale voltage is predicted to be 2.5 mV for a 0.1 % strain. If we assume a $1/f$ noise corner frequency of 5 Hz, $f_l = 0.01$ Hz, and a first-order bandwidth of 1 kHz ($k_e = 1.57$). The resolution predicted by Eq. (8.29) is 580 nV or 230 ppm. In other words, if the full-scale range was $100\,\mu\text{m}$, the resolution would be 23 nm, which is not competitive.

8.3.2 Piezoresistive Strain Sensors

In 1954, a visiting researcher at Bell Laboratories, C. S. Smith, demonstrated that "exceptionally large" resistance changes occur in silicon and germanium when subjected to external strain [33]. This discovery was the foundation for today's semiconductor piezoresistive sensors that are now ubiquitous in applications such as integrated pressure sensors and accelerometers [34].

Compared to metal foil strain gauges that respond only to changes in geometry, piezoresistive sensors exhibit up to two orders-of-magnitude greater sensitivity. In addition to their high strain sensitivity, piezoresistive sensors are also easily integrated into standard integrated circuit and MEMS fabrication processes which is highly advantageous for both size and cost. The foremost disadvantages associated with piezoresistive sensors are the low strain range (0.1 %), high temperature sensitivity, poor long-term stability, and slight nonlinearity (1 %) [34]. The elimination of these artifacts requires a more complicated conditioning circuit than metal foil strain gauges; however, integrated circuits are now available that partially compensate for nonlinearity, offset, and temperature dependence, for example, the Maxim MAX1450.

As shown in Fig. 8.10, a typical integrated piezoresistive strain sensor consists of a planar n-doped resistor with heavily doped contacts. When the sensor is elongated in the x-axis, the average electron mobility increases in that direction, reducing resistance [34]. The effect is reverse during compression, or if the resistor is p-type. Since the piezoresistive effect is due to changes in the crystal lattice, the effect is highly dependent on the crystal orientation. The change in resistance can be expressed as,

$$\Delta R = R_G \left[\pi_L \sigma_{xx} + \pi_T \left(\sigma_{yy} + \sigma_{zz} \right) \right], \tag{8.38}$$

where ΔR is the change in resistance; R_G is the nominal resistance; σ_{xx}, σ_{yy}, and σ_{zz} are the tensile stress components in each axis; and π_L and π_T are the longitudinal and transverse piezoresistive coefficients which are determined from the crystal orientation [34].

Due to the temperature dependence and low strain range, piezoresistive sensors are primarily used in microfabricated devices where the difficulties are offset by the high sensitivity and ease of fabrication, for example, meso-scale nanopositioners

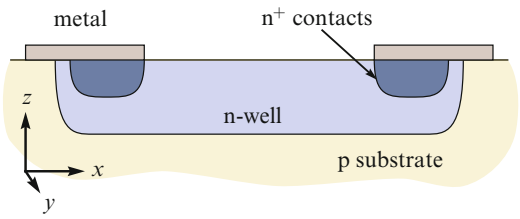


Fig. 8.10 A cross-section of a piezoresistive strain sensor. Deformation of the semiconductor crystal causes a resistance change one-hundred times that of a resistive strain gauge

[35] and MEMs devices [36]. Discrete piezoresistive sensors are also available for standard macro-scale nanopositioning applications, for example, Micron Instruments SS-095-060-350PU. Discrete piezoresistive strain sensors are significantly smaller than metal foil gauges, for example, the Micron Instruments SS-095-060-350PU is 2.4 mm×0.4 mm. The sensitivity is typically specified in the same way as a metal foil sensor, by the gauge factor defined in Eq. (8.33). While the gauge factor of a metal foil sensor is between 1 and 2, the gauge factor of the Micron Instruments SS-095-060-350PU is 120.

Due to the temperature dependence of piezoresistive strain sensors, practical application requires a closely collocated half- or full-bridge configuration, similar to a metal foil gauge. The required signal conditioning is also similar to the metal foil gauges. If an accuracy of better than 1 % is required, or if large changes in temperature are expected, the piezoresistive elements must be closely matched and the signal conditioning circuit must be compensated for temperature and nonlinearity. Two fully integrated bridge conditioning circuits include the MAX1450 and MAX1452 from Maxim Integrated Products, USA.

Alike metal foil strain gauges, the noise in piezoresistive sensors is predominantly thermal and $1/f$ noise [34]. However, since piezoresistive sensors are semiconductors, the $1/f$ noise can be substantially worse [34]. Consider the Micron Instruments SS-095-060-350PU piezoresistive sensor which has a gauge factor of 120 and a resistance of 350 Ω. In a two-varying-element bridge with 2-V excitation, Eq. (8.36) predicts that a full-scale strain of 0.1 % develops 120 mV. The thermal noise due to the resistance is 2.4 nV/$\sqrt{\text{Hz}}$. If the $1/f$ noise corner frequency is assumed to be 10 Hz, the resolution with a first-order bandwidth of 1000 Hz is 130 nV which implies a 6σ-resolution of 590 nV or 4.9 ppm. Restated, if the full-scale displacement was 100 μm, the resolution would be 0.49 nm.

Although the majority of piezoresistive sensors are integrated directly into MEMS devices, discrete piezoresistive strain sensors are available from: Kulite Semiconductor Products Inc., USA; and Micron Instruments, USA.

8.3.3 Piezoelectric Strain Sensors

In addition to their actuating role, piezoelectric transducers are also widely utilized as high sensitivity strain sensors [37–43].

The basic operation of a piezoelectric strain sensor is illustrated in Fig. 8.11a. In this case the applied force F and resulting strain $\Delta h/h$ is aligned in the same axis as the polarization vector. The polarization vector points in the same direction as the internal dipoles which is opposite in direction to the applied electric field. Thus, compression of the actuator results in a voltage of the same polarity as the voltage applied during polarization. From the stress-voltage form of the piezoelectric constituent equations, the developed electric field E is

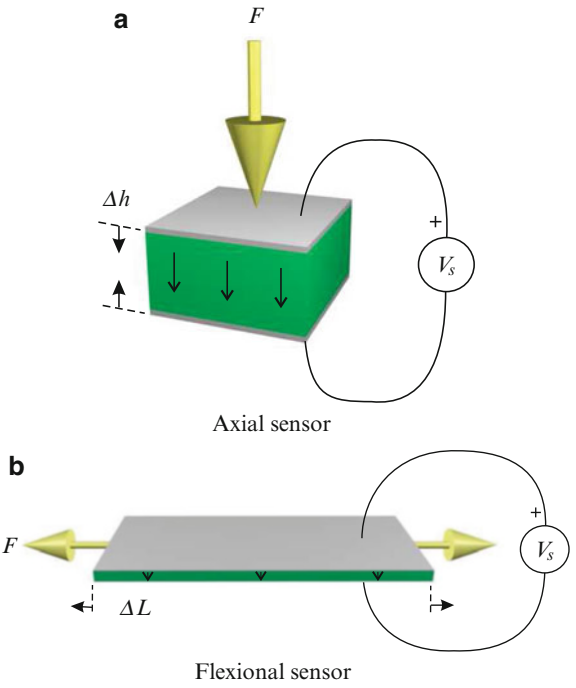


Fig. 8.11 A piezoelectric stack and plate strain sensor. The polarization vector is shown as a *downward arrow*. (**a**) Axial sensors are typically used to measure dynamic forces while (**b**) flexional sensors are used to measure changes in strain or curvature

$$E = q_{33}\frac{\Delta h}{h}, \tag{8.39}$$

where Δh is the change in thickness, h is the thickness, and q_{33} is the piezoelectric coupling coefficient for the stress-voltage form. The constant q_{33} is related to the piezoelectric strain constant d_{33} by

$$q_{33} = \frac{d_{33}}{\epsilon^T s^D}, \tag{8.40}$$

where ϵ^T is the permittivity under constant stress (in Farad/m) and s^D is the elastic compliance under constant electric displacement (in m^2/N). If the piezoelectric voltage constant g_{33} is known instead of q_{33} or d_{33}, q_{33} can also be derived from $q_{33} = g_{33}/s^D$. By multiplying (8.40) by the thickness h, the measured voltage can be written as

$$V_s = q_{33}\Delta h, \tag{8.41}$$

If there are multiple layers, the voltage is

$$V_s = \frac{q_{33}}{n}\Delta h, \tag{8.42}$$

where n is the number of layers. The developed voltage can also be related to the applied force [31].

$$V_s = \frac{nd_{33}}{C}F, \quad \text{or} \quad V_s = \frac{d_{33}h}{n\epsilon^T A}F, \tag{8.43}$$

where C is the transducer capacitance defined by $C{=}n^2\epsilon^T A/h$ and A is the area.

The voltage developed by the flexional sensor in Fig. 8.11b is similar to the axial sensor except for the change of piezoelectric constant. In a flexional sensor, the applied force and resulting strain are perpendicular to the polarization vector. Hence, the g_{31} constant is used in place of the g_{33} constant. Assuming that the length L is much larger than the width and thickness, the developed voltage is

$$V_s = \frac{-g_{31}}{L}F, \tag{8.44}$$

which can be rewritten in terms of the stiffness k and strain,

$$V_s = -g_{31}k\frac{\Delta L}{L} \tag{8.45}$$

$$V_s = \frac{-g_{31}A}{s^D L}\frac{\Delta L}{L}, \tag{8.46}$$

where A is the cross-sectional area equal to width $\times$ thickness.

When mounted on a host structure, flexional sensors can be used to detect the underlying stress or strain as well as the curvature or moment [37, 44, 45]. In nanopositioning applications, the electrodes of a piezoelectric tube act as a plate sensor and can be used to detect the strain and hence displacement [39, 40, 42]. This application is illustrated in Fig. 8.12.

Due to the high mechanical stiffness of piezoelectric sensors, thermal or Boltzmann noise is negligible compared to the electrical noise arising from interface electronics. As piezoelectric sensors have a capacitive source impedance, the noise density $N_{V_s}(\omega)$ of the sensor voltage V_s is due primarily to the current noise i_n generated by the interface electronics. The equivalent electrical circuit of a piezoelectric sensor and high-impedance buffer is shown in Fig. 8.13. Neglecting the leakage resistance R, the noise density of the sensor voltage is

$$N_{V_s}(\omega) = i_n\frac{1}{C\omega}, \tag{8.47}$$

where N_{V_s} and i_n are the noise densities of the sensor voltage and current noise, measured in Volts and Amps per $\sqrt{\text{Hz}}$, respectively.

The experimentally measured and predicted noise density of a piezoelectric sensor is plotted in Fig. 8.14. The sensor is a 2-mm Noliac CMAP06 stack mounted on top of 10-mm long actuator, the assembly is mounted in the nanopositioning stage pictured in Fig. 8.15. The sensor has a capacitance of 30 nF and the voltage

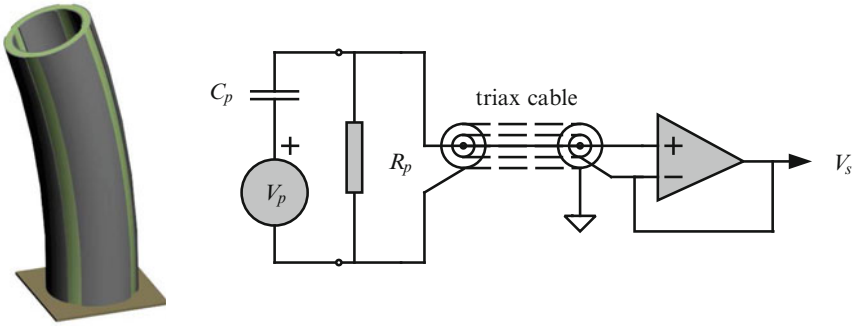


Fig. 8.12 A piezoelectric tube actuator with one electrode utilized as a strain sensor. The electrical equivalent circuit consists of the induced piezoelectric voltage V_p in series with the transducer capacitance. The dielectric leakage and input impedance of the buffer circuit are modeled by the parallel resistance R_p. An effective method for shielding the signal is to use a triaxial cable with the intermediate shield driven at the same potential as the measured voltage. (Tube drawing courtesy K. K. Leang)

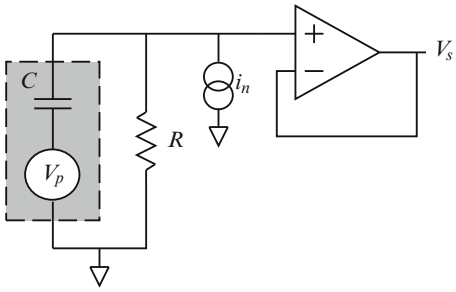
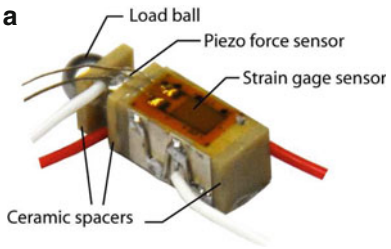


Fig. 8.13 The electrical model of a piezoelectric force sensor. The open-circuit voltage V_p is high-pass filtered by the transducer capacitance C and leakage resistance R. The current source i_n represents the current noise of a high-impedance buffer

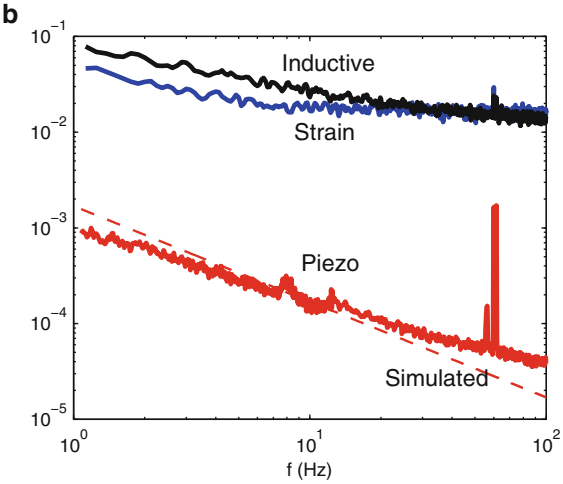
buffer (OPA606) has a noise density of 2 fA/$\sqrt{\text{Hz}}$. Further details on the behavior of piezoelectric force sensors can be found in [31].

In Fig. 8.14b the noise density of the piezoelectric sensor is observed to be more than two orders of magnitude less than the strain and inductive sensors at 100 Hz. The noise density also continues to reduce at higher frequencies. However, at low frequencies the noise of the piezoelectric force sensor eventually surpass the other sensors. As the noise density is equivalent to an integrator excited by white noise, the measured voltage drifts significantly at low frequencies. A time record that illustrates this behavior is plotted in Fig. 8.16. The large drift amplitude is evident. Thus, although the piezoelectric force sensor generates less noise than the strain and inductive sensors at frequencies in the Hz range and above, it is inferior at frequencies below approximately 0.1 Hz.

In addition to noise, piezoelectric force sensors are also limited by dielectric leakage and finite buffer impedance at low frequencies. The induced voltage V_p shown in Fig. 8.13 is high-pass filtered by the internal transducer capacitance C and the leakage resistance R. The cut-off frequency is



Actuator with integrated force sensor and strain gauge



Noise power spectral density (nm/$\sqrt{\text{Hz}}$)

Fig. 8.14 (**a**) A piezoelectric stack actuator with an integrated force sensor and two resistive strain gauges bonded to the top and bottom surface (the bottom gauge is not visible). In (**b**), the noise density of the piezoelectric sensor is compared to the resistive strain gauge and a Kaman SMU9000-15N inductive sensor, all signals are scaled to nm/$\sqrt{\text{Hz}}$. The simulated noise of the piezoelectric force sensor is also plotted as a *dashed line*

$$f_{hp} = \frac{1}{2\pi RC} \text{ Hz}. \tag{8.48}$$

The buffer circuit used in the results above has an input impedance of 100 MΩ, this results in a low-frequency cut-off of 0.05 Hz. To avoid a phase lead of more than 6 degrees, the piezoelectric force sensor cannot be used to measure frequencies of less than 0.5 Hz.

Piezoelectric actuators and sensors are commercially available from: American Piezo (APC International, Ltd.), USA; CeramTec GmbH, Germany; Noliac A/S, Denmark; Physik Instrumente (PI), Germany; Piezo Systems Inc., USA; and Sensor Technology Ltd., Canada.

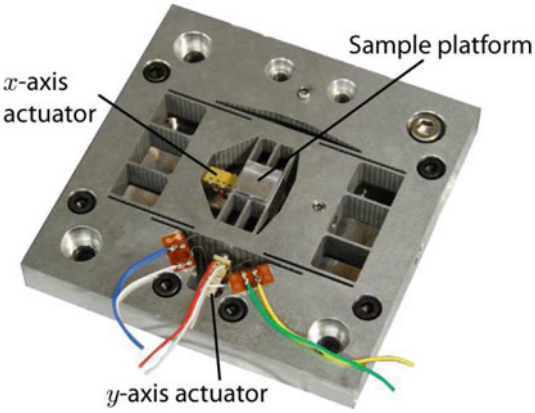


Fig. 8.15 A nanopositioning platform with a two-varying-element strain gauge fitted to the y-axis actuator [31]. The nanopositioner is driven by two piezoelectric stack actuators that deflect the sample platform by a maximum of 10 μm in the x and y lateral axes

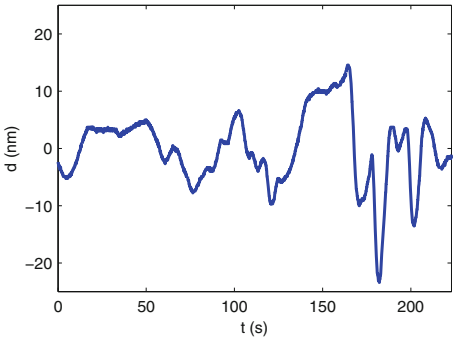


Fig. 8.16 Low-frequency noise of the piezoelectric sensor pictured in Fig. 8.14a, scaled to nanometers. The peak-to-peak noise over 220 s is 38 nm or 26 mV

8.3.4 Capacitive Sensors

Capacitive sensors are the most commonly used sensors in short-range nanopositioning applications. They are relatively low-cost and can provide excellent linearity, resolution, and bandwidth [46]. However, due to the electronics required for measuring the capacitance and deriving position, capacitive sensors are inherently more complex than sensors such as resistive strain gauges. Larger ranges can be achieved with the use of an encoder-style electrode array [47].

All capacitive sensors work on the principle that displacement is proportional to the change in capacitance between two conducting surfaces. If fringe effects are neglected, the capacitance C between two parallel surfaces is

$$C = \frac{\epsilon_0 \epsilon_r A}{h}, \tag{8.49}$$

where ϵ_0 is the permittivity of free space, ϵ_r is the relative permittivity of the dielectric (or dielectric constant), A is the area between the surfaces, and h is the distance between the surfaces.

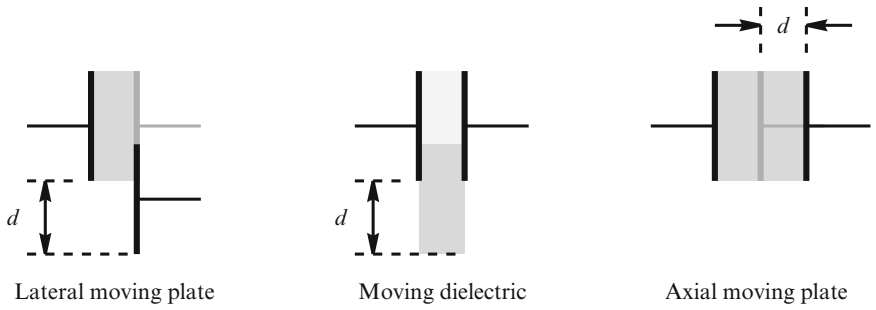


Fig. 8.17 Types of capacitive sensor. The axial moving plate produces the highest sensitivity but the smallest practical travel range. Lateral moving plate and moving dielectric sensors are most useful in long-range applications

Three types of capacitive sensor are illustrated in Fig. 8.17. The lateral moving plate design is used for long-range measurements where the plate spacing can be held constant. This is often achieved with two concentric cylinders mounted on the same axis. In this configuration, the change in capacitance is proportional to the change in area and hence position. A similar arrangement can be found in the moving dielectric sensor where the area and distance are constant but the dielectric is variable. This approach is not commonly used because a solid dielectric is required that causes friction and mechanical loading.

The axial moving plate, or parallel plate capacitive sensor is the most common type used in nanopositioning applications. Although the useful range is smaller than other configurations, the sensitivity is proportionally greater. The capacitance of a moving plate sensor is

$$C = \frac{\epsilon_0 \epsilon_r A}{d}, \tag{8.50}$$

hence, the sensitivity is

$$\frac{\mathrm{d}\, C}{\mathrm{d}\, d} = \frac{C_0}{d_0}\ \mathrm{F/m}, \tag{8.51}$$

where C_0 and d_0 are the nominal capacitance and distance. Thus, for a sensor with a nominal capacitance of 10 pF and spacing of 100 μm, the sensitivity is 100 fF/μm. The sensitivity of different capacitive sensor types is compared in [20].

A practical parallel plate capacitive sensor is illustrated in Fig. 8.18. In addition to the probe electrode, a guard electrode is also used to shield the probe from nearby electric fields and to improve linearity. The guard electrode is driven at the same potential as the probe but is not included in the capacitance measurement. As the fringing effect in the electric field is only present at the outside electrode, the nonlinearity in the capacitance measurement and distance calculation is reduced. A summary of correction terms for different guard electrode geometries can be found in references [20] and [46].

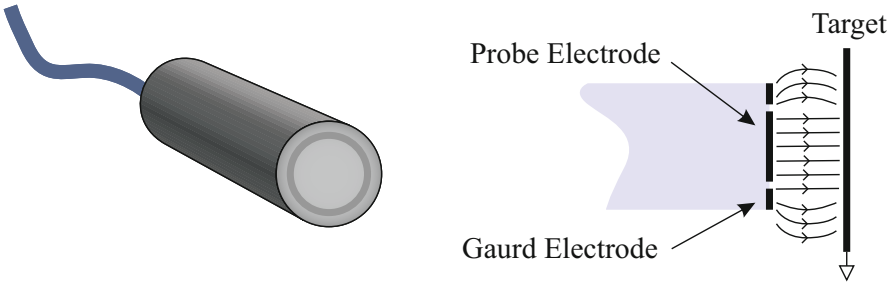


Fig. 8.18 A capacitive sensor probe and electrode configuration. The guard electrode is driven at the same potential as the probe in order to linearize the electric field and reduce fringing effects

To measure the capacitance and thus derive the position, a wide variety of circuits are available [21, 46]. The simplest circuits are timing circuits where the timing capacitor is replaced by the sensor capacitance. Examples include the ubiquitous 555 timer in the one-shot or free-running oscillator modes. The output of a one-shot circuit is a pulse delay proportional to the capacitance. Likewise, the output of the oscillator is a square-wave whose frequency is proportional to capacitance. Although these techniques are not optimal for nanopositioning applications, they are simple, low-cost, and can be directly connected to a microcontroller with no analog-to-digital converters.

A direct measurement of the capacitance can be obtained by applying an AC voltage V to the probe electrode and grounding the target. The resulting current I is determined by Ohms law,

$$I = j\omega VC, \tag{8.52}$$

where ω is the excitation frequency in rad/s. Since the current is proportional to capacitance, this method is useful for the lateral moving plate and moving dielectric configurations where the displacement is also proportional to capacitance. For the axial moving plate configuration, where the displacement is inversely proportional to capacitance, it is more convenient to apply a current and measure the voltage. In this case, the measured voltage in response to an applied current is

$$V = \frac{I}{j\omega C}, \tag{8.53}$$

which is inversely proportional to capacitance and thus proportional to displacement.

Regardless of whether the current or voltage is the measured variable, it is necessary to compute the AC magnitude of the signal. The simplest circuit that achieves this is the single-diode demodulator or envelope detector shown in Fig. 8.19a. Although simple, the linearity and offset voltage of this circuit are dependent on the diode characteristics which are highly influenced by temperature.

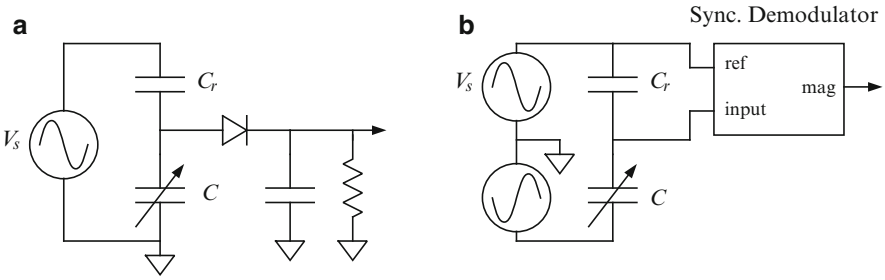


Fig. 8.19 Demodulation circuits for measuring capacitance. The linearity, temperature sensitivity, and noise performance of the synchronous detector are significantly better than the single-diode envelope detector. (**a**) single-diode demodulator and (**b**) Synchronous demodulator

A better option is the synchronous demodulator with balanced excitation shown in Fig. 8.19b. A synchronous demodulator can be constructed from a filter and voltage controlled switch [21, 46]. Integrated circuit demodulators such as the Analog Devices AD630 are also available. Synchronous demodulators provide greatly improved linearity and stability compared to single-diode detectors.

The balanced excitation in Fig. 8.19b eliminates the large DC offset produced by single ended demodulators, such as Fig. 8.19a. The balanced configuration also eliminates the offset sensitivity to changes in the supply voltage, which greatly improves the stability. Although single ended excitation can be improved with a full-bridge configuration, this requires a high common-mode rejection ratio, which is difficult to obtain at high frequencies.

In general, capacitive sensors with guard electrodes can provide excellent linearity in ideal conditions (10 ppm or 0.001 %); however, practical limitations can significantly degrade this performance. A detailed analysis of capacitive sensor nonlinearity in [20] concluded that the worst sources of nonlinearity are tilting and bowing. Tilting is the angle between the two parallel plates and bowing is the depth of concavity or convexity.

A summary of the error analysis performed in [20] is contained in Table 8.2. Considering that the linearity of an capacitive sensor in ideal conditions can be 0.001 %, the effect of tilting and bowing severely degrades the performance. These errors can be reduced by careful attention to the mounting of capacitance sensors. It is recommended that capacitive sensors be fixed with a spring washer rather than a screw. This can significantly reduce mounting stress on the host structure and sensor. In addition to deformation, excessive mounting forces can slowly relieve over time causing major drifts in offset, linearity, and sensitivity.

The magnitude of error due to tilting and bowing can be reduced by increasing the nominal separation of the two plates, this also increases the range. However, if the area of the sensor is not increased, the capacitance drops, which increases noise.

The noise developed by a capacitive sensor is due primarily to the thermal and shot-noise of the instrumentation electronics. Due to the demodulation process, the noise spectral density is relatively flat and does not contain a significant

Table 8.2 A summary of error sources in a parallel plate capacitive sensor studied in [20]

Errors due to tilting		
Tilt angle (mrad)	2	5
Nonlinearity (%)	0.08	0.6
Offset (%)	0.35	2.4
Scale error (%)	0.8	5.4
Errors due to bowing		
Bow depth (μm)	10	30
Nonlinearity (%)	0.025	0.33
Offset (%)	5	18
Scale error (%)	3	11

The sensor has a gap of 100 μm, a radius of 6 mm, and a nominal capacitance of 10 pF

$1/f$ component. Although the electronic noise remains constant with different sensor configurations, the effective position noise is proportional to the inverse of sensitivity. As the sensitivity is C_0/d_0 (8.51), if the capacitance is doubled by increasing the area, the position noise density is reduced by half. However, if the nominal gap d_0 is doubled to improve the linearity, the capacitance also halves, which reduces the sensitivity and increases the noise density by a factor of four. The position noise density is minimized by using the smallest possible plate separation and the largest area.

A typical commercial capacitive sensor with a range of 100 μm has a noise density of approximately of 20 pm/$\sqrt{\text{Hz}}$ [40]. The $1/f$ corner frequency of a capacitive sensor is typically very low, around 10 Hz. With a first-order bandwidth of 1 kHz, the resolution predicted by Eq. (8.29) is 2.4 nm or 24 ppm. This can be reduced to 0.55 nm or 5.5 ppm by restricting the bandwidth to 10 Hz.

Capacitive position sensors are commercially available from: Capacitec, USA; Lion Precision, USA; Micro-Epsilon, Germany; MicroSense, USA; Physik Instrumente (PI), Germany; and Queensgate Instruments, UK. Two commercially available devices are pictured in Fig. 8.20.

8.3.5 MEMs Capacitive and Thermal Sensors

MEMs capacitive sensors operate on a similar principles to their macro-scale counterpart discussed in the previous section. However, due to their small size, a more complicated geometry is required to achieve a practical value of capacitance. The comb-type sensor illustrated in Fig. 8.21a is a common variety found in a number of nanopositioning applications, for example, [48, 49]. In this configuration, the total capacitance is approximately proportional to the overlap area of each electrode array.

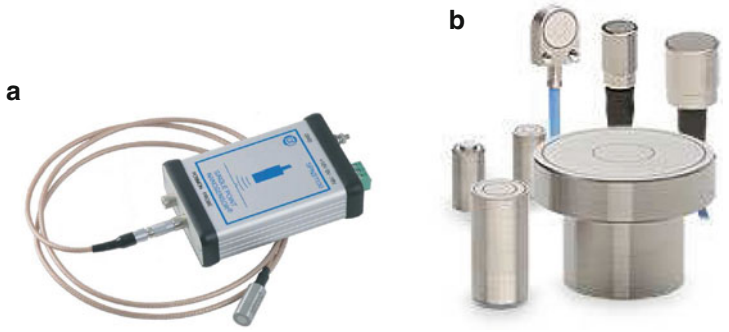


Fig. 8.20 An example of two commercially available capacitive sensors. Photos courtesy of (**a**) Queensgate Instruments, UK and (**b**) Micro-Epsilon, Germany

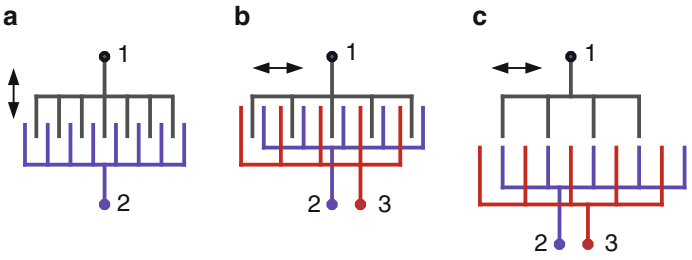


Fig. 8.21 Three examples of MEMs capacitive sensor geometries. (**a**) Standard comb sensor; (**b**) Differential comb sensor; and (**c**) Incremental capacitive encoder

The basic comb sensor can be improved by employing a differential detection method as illustrated in Fig. 8.21b. Here, two sets of excitation electrodes (terminals 2 and 3) are driven 180 degrees out of phase. Thus, at the central position, the potential at terminal 1 is zero. This configuration provides a higher sensitivity than the basic comb sensor and is used extensively in devices such as accelerometers and gyroscopes [46, 50].

To increase the range of motion beyond a single inter-electrode spacing, the configuration in Fig. 8.21c uses withdrawn electrodes to form a capacitive incremental encoder [51–53]. The slider can now move freely in either direction, limited only by the length of the excitation array. As the slider moves horizontally, the induced voltage at terminal 1 alternates between the phase of terminals 2 and 3. A second array is typically used to create a quadrature signal for ascertaining the direction of travel. This approach can provide a large travel range with high resolution but the decoding electronics is more complicated and the performance is sensitive to the separation between the arrays. If the two arrays can be overlain vertically, the capacitance can be increased while the difficulties with array separation are reduced [54, 55].

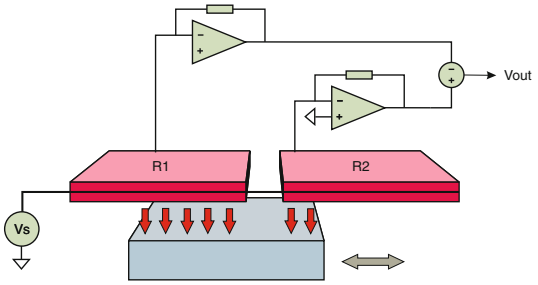


Fig. 8.22 An electrothermal position sensor. The two stationary microheaters are driven by a constant voltage source Vs. The rate of heat transfer and the resulting temperature is proportional to the overlap between the heater and the heatsink. The position of the heatsink can be estimated by measuring the current difference between the two microheaters which indicates the difference in resistance and temperature

Electrothermal sensors are an alternate class of position sensors first utilized in nanopositioning applications by IBM in 2005 [56]. An example of a differential electrothermal position sensor is illustrated in Fig. 8.22. Two microheaters are driven by a DC voltage source resulting in a temperature increase. Due to the heat transfer between the microheater and moving heatsink, the temperature of each microheater becomes a function of the overlap area and hence position. The heatsink position is estimated by measuring the difference in current which is related to the resistance and temperature.

An advantage of electrothermal sensors over capacitive sensors is the compact size which has made them appealing in applications such as data storage [57–59] and nanopositioning [60, 61]. The noise performance of electrothermal sensors can be similar or superior to capacitive sensors under certain conditions. However, due to the elevated temperature, electrothermal sensors are known to exhibit a significant amplitude of low-frequency noise [49].

With a range of 100 μm, a thermal position sensing scheme achieved a noise density of approximately 10 pm/$\sqrt{\text{Hz}}$ with a 1/f corner frequency of approximately 3 kHz [60]. This resulted in a resolution of 10 nm over a bandwidth of 4 kHz. As a result of the low-frequency noise and drift, an auxiliary position sensor was utilized at frequencies below 24 Hz [60].

8.3.6 Eddy-Current Sensors

Eddy-current, or inductive proximity sensors, operate on the principle of electromagnetic induction [62, 63]. As illustrated in Fig. 8.23, an eddy-current probe consists of a coil facing an electrically conductive target. When the coil is excited by an AC current, the resulting magnetic field passes through the conductive target and induces a current according to Lenz's law. The current flows at right angles to

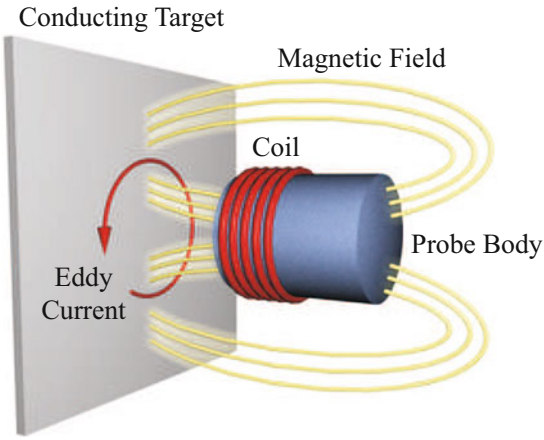


Fig. 8.23 The operating principle of an eddy-current sensor. An alternating current in the coil induces eddy currents in the target. Increasing the distance between the probe and target reduces the eddy currents and also the effective resistance of the coil

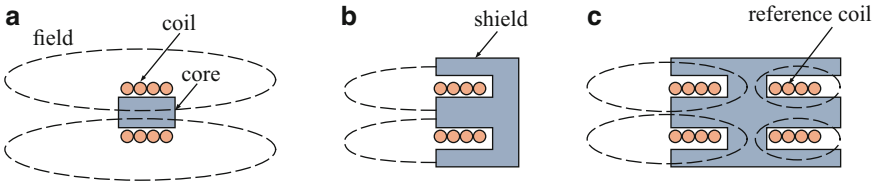


Fig. 8.24 Types of eddy-current sensor. The unshielded type has the greatest range but is affected by nearby fields and conductors. A shield makes the magnetic field more directional but reduces the range. A reference coil can be used to reduce the sensitivity to temperature.(**a**) Unshielded, (**b**) Shielded, and (**c**) Balanced

the applied magnetic field and develops an opposing field. The eddy currents and opposing field become stronger as the probe approaches the target.

The distance between probe and target is detected by measuring the AC resistance of the excitation coil which depends on the magnitude of the opposing field and eddy current. The required electronics are similar to that of a capacitive sensor and include an oscillator and demodulator to derive the resistance [21, 62, 64].

Three common types of eddy-current sensor are depicted in Fig. 8.24. The unshielded sensor has a large magnetic field that provides the greatest range; however, it also requires the largest target area and is sensitive to nearby conductors. Shielded sensors have a core of permeable material such as Permalloy which reduces the sensitivity to nearby conductors and requires less target area; however, they also have less range. The balanced type has a second shielded or non-inductive coil that is used to null the effect of temperature variation [65]. The second coil is used in a divider or bridge configuration such as that illustrated in Fig. 8.25.

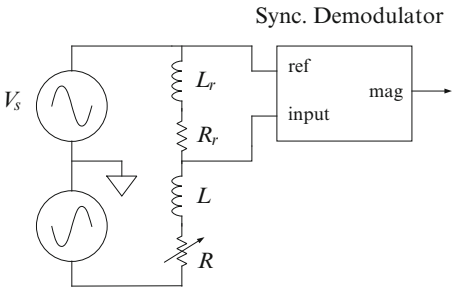


Fig. 8.25 Synchronous demodulation circuit for a balanced eddy-current sensor. L_r and R_r are the inductance and resistance of the reference coil

Another type of position sensor similar to an eddy-current sensor is the inductive proximity sensor, also referred to as a differential reluctance transducer if a reference coil is present. Rather than a conductive target, an inductive proximity sensor requires a ferromagnetic target. Since the reluctance of the magnetic path is proportional to the distance between the probe and target, the displacement can be derived from the coil inductance. Inductive proximity sensors have the same construction and electronics requirement as an eddy-current sensor. Their main drawback compared to eddy-current sensors is the temperature dependent permeability of the target material and the presence of magnetic hysteresis.

Eddy-current sensors are not as widely used as capacitive sensors in nanopositioning applications due to the temperature sensitivity and range concerns. The temperature sensitivity arises from the need of an electrical coil in the sensor head and the varying resistance of the target. The minimum range of an eddy-current sensor is limited by the minimum physical size of the coil, which imposes a minimum practical range of between 100 and 500 μm. In contrast, capacitive sensors are available with a range of 10 μm which can provide significantly higher resolution in applications with small travel ranges.

The major advantage of eddy-current and inductive sensors is the insensitivity to dust and pollutants in the air-gap and on the surface of the sensor. This gives them a significant advantage over capacitive sensors in industrial applications.

The noise performance of an eddy-current sensor can be similar to that of a capacitive sensor. For example, the noise density of the Kaman SMU9000-15N which has a range of 500 μm is plotted in Fig. 8.14b. The $1/f$ corner frequency is approximately 20 Hz and the constant density is approximately 20 pm/$\sqrt{\text{Hz}}$. Equation (8.29) predicts a resolution of 5 nm or 10 ppm with a bandwidth of 1 kHz. Due to the physical size of the coils, smaller ranges and higher resolution is difficult to achieve.

Eddy-current position sensors are commercially available with ranges of approximately 100 μm to 80 mm. Manufacturers include: Micro-Epsilon, Germany; Kaman Sensors, USA; MicroStrain, USA; Keyence, USA; Lion Precision, USA; and Ixthus Instrumentation, UK. Two commercially available devices are pictured in Fig. 8.26.

Fig. 8.26 Two commercially available eddy-current sensors. (**a**) Photos courtesy of Lion Precision, USA and (**b**) Micro-Epsilon, Germany

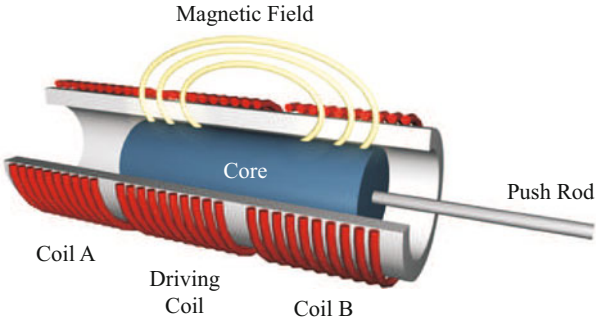


Fig. 8.27 The operating principle of a linear variable displacement transducer (LVDT). Changes in the core position produce a linear differential change in the coupling between the driving coil and the pick-up coils

8.3.7 Linear Variable Displacement Transformers

Linear variable displacement transformers (LVDTs) are used extensively for displacement measurement with ranges of 1 mm to over 50 cm. They were originally described in a patent by G. B. Hoadley in 1940 (US Patent 2,196,809) and became popular in military and industrial applications due to their ruggedness and high resolution [21].

The operating principle of an LVDT is illustrated in Fig. 8.27. The stationary part of the sensor consists of a single driving coil and two sensing coils wound onto a thermally stable bobbin. The movable component of the transducer is a permeable material such as Nickel–Iron (Permalloy), and is placed inside the bobbin. The core is long enough to fully cover the length of at least two coils. Thus, at either extreme, the central coil always has a complete core at its center.

Since the central coil always has a complete core, all of the magnetic flux is concentrated in the core. As the core moves, the amount of flux passing through each sensor coil is proportional to the length of core contained within. Hence, the

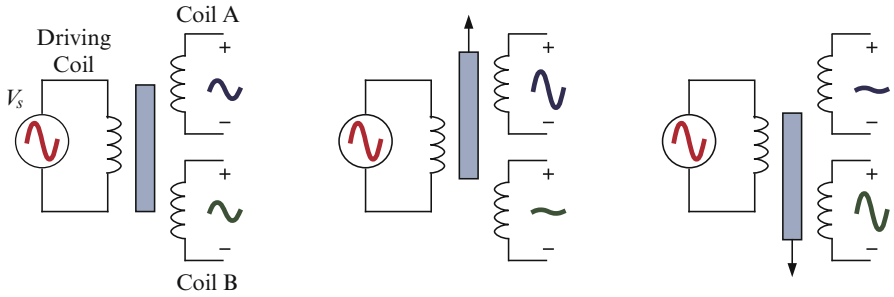


Fig. 8.28 The relationship between the sensor coil voltage and core position in an LVDT. The coil voltage is proportional to the amount or core it contains

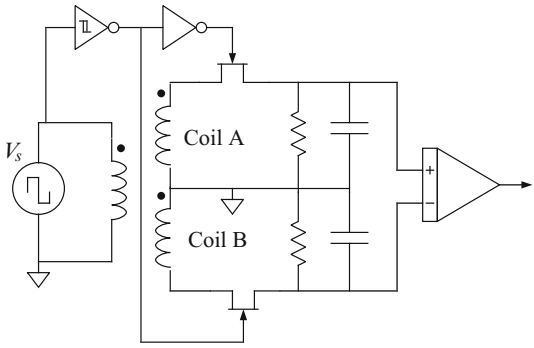


Fig. 8.29 An LVDT conditioning circuit with a synchronous demodulator and differential amplifier [21]

displacement of the core is proportional to the difference in voltage induced in the sensor coils. This principle is illustrated in Fig. 8.28.

In addition the components in Fig. 8.27, a bearing is required to guide the motion of the core through the bobbin. An external case is also required that can be constructed from a permeable material to provide magnetic shielding of the coils. It is important that the push-rod be constructed from a non-magnetic material such as aluminum or plastic, otherwise it contributes erroneously to the coupling between the coils.

The electronics required by an LVDT are similar to that required for a capacitive or inductive sensor. An oscillator excites the driving coil with a frequency of around 1 kHz. Although higher frequencies increase the sensor bandwidth they also induce eddy currents in the core that are detrimental to performance [21]. Alike a capacitive or eddy-current sensor, a demodulator is required to determine the AC magnitude of the voltage induced in each coil. A simple synchronous demodulator circuit for this purpose is shown in Fig. 8.29 [21]. The square-wave oscillator is replaced by a sine-wave oscillator if the electronics and LVDT are not physically collocated. Other demodulation circuits include the single-diode demodulator in Fig. 8.19a and the AD630-based demodulator in Fig. 8.19b.

The greatest advantages of LVDTs are the infinitesimal resolution, large range, simplicity, and ruggedness. Very low levels of electrical noise can be achieved due to the low-impedance of the sensing coils. Nonlinearity is also below 1 % without the need for field calibration or mapping functions. The major drawbacks of LVDTs include the limited bandwidth and sensitivity to lateral motion. Due to eddy currents and the inter winding capacitance, the excitation frequency is limited to a few tens of kHz, which limits the bandwidth to between 100 Hz and 1 kHz. Although classified as a non-contact sensor, bearings are required to guide the core linearly through the bobbin. This can be a significant disadvantage in nanopositioning applications if the sensor adds both friction and mass to the moving platform. However, if the platform is already flexure-guided, additional bearings may not be required. LVDTs are most suited to one-degree-of-freedom applications with relatively large displacement ranges of approximately 1 mm or greater. A range of less than 0.5 mm is difficult to achieve due to the small physical size of the coils. A notable exception is the air core LVDT coils used to detect position in the Asylum Research (USA) atomic force microscopes [66]. The air core eliminates eddy-current losses and Barkhausen noise caused by the high permeability materials. An RMS noise of 0.19 nm was reported for a range of 16 μm which equates to a resolution of approximately 1.14 nm and a dynamic range of 71 ppm [66].

The theoretical resolution of LVDT sensors is limited primarily by the Johnson noise of the coils and Barkhausen noise in the magnetic materials [66]. However, standard conditioning circuits like the Analog Devices AD598 produce electronic noise on the order of 50 μVp-p with a bandwidth of 1 kHz. This imposes a resolution of approximately 10 ppm when using a driving amplitude of 5 Vp-p. Since the smallest commercially available range is 0.5 mm, the maximum resolution is approximately 5 nm with a 1 kHz bandwidth.

Due to their popularity, LVDTs and the associated conditioning electronics are widely available. Some manufacturers of devices that may be suitable in micro- and nanopositioning applications include: Macro Sensors, USA; Monitran, UK; Singer Instruments, Israel; MicroStrain, USA; Micro-Epsilon, USA; and Honeywell, USA. Two commercially available LVDTs are pictured in Fig. 8.30.

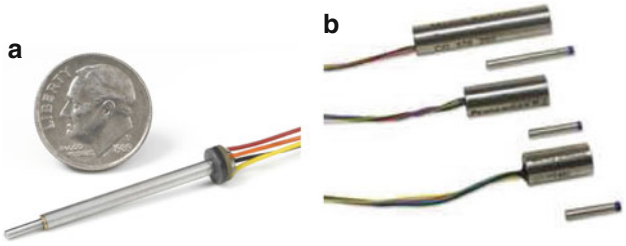


Fig. 8.30 Commercially available LVDT sensors. Photos courtesy of (**a**) Singer Instruments, Israel and (**b**) Macro Sensors, USA

8.3.8 *Laser Interferometers*

Since 1960, the meter length standard has been defined by optical means. This change arose after Michelson invented the interferometer which improved the accuracy of length measurement from a few parts in 10^7, to a few parts in 10^9 [67]. Thus, in 1960, the meter was redefined in terms of the orange line from a ^{86}Kr discharge lamp.

In 1983, the meter was redefined as the length traveled by light in a vacuum during a time interval of 1/299 792 458 s [67]. This definition was chosen because the speed of light is now fixed and the primary time standard, based on the ^{133}Cs clock, is known to an accuracy of a few parts in 10^{11} [67]. Length measurements are performed by interferometry using lasers with a frequency measured against the time standard. With a known frequency and speed, the laser wavelength can be found to an extremely high accuracy. Stabilized lasers are now available with precisely calibrated wavelengths for metrological purposes. Metrological traceability is described further in Sect. 8.2.7.

The operating principle of a Michelson interferometer is described in Fig. 8.31. A laser beam is split into two paths, one that is reflected by a moving mirror and another reflected by a stationary mirror. The movement of the mirror is measurable by observing the fringe pattern and intensity at the detector. If the distance between the paths is an integer number of wavelengths, constructive interference occurs. The displacement of the moving mirror, in wavelengths, is measured by counting the number of interference events that occur. The phase of the interference, and hence the displacement between interference events, can also be derived from the detector intensity.

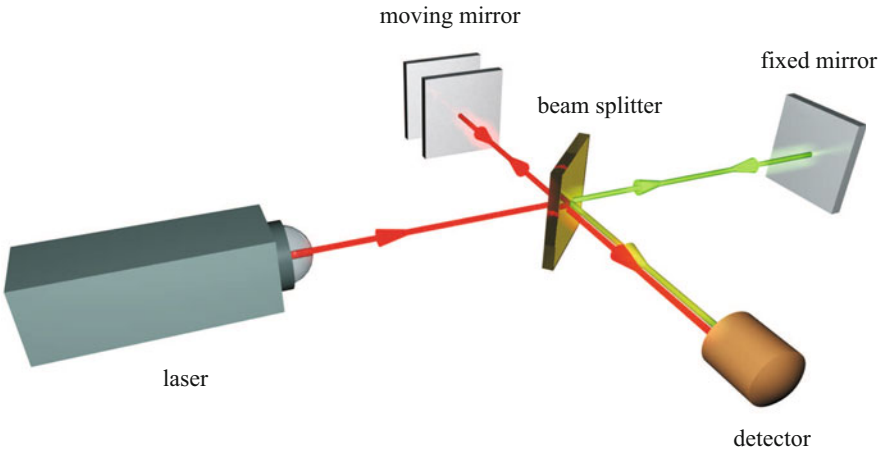


Fig. 8.31 The operation of a Michelson interferometer. The laser light is split into two paths, one that encounters a moving mirror and another that is fixed. The two beams are recombined and interfere at the detector. If the distance between the paths is an integer number of wavelengths, constructive interference occurs

Although simple, the Michelson interferometer is rarely used directly for displacement metrology. Due to the reference path, the Michelson interferometer is sensitive to changes or movement in the reference mirror and the beam splitter. Differences between the optical medium in the reference and measurement path are also problematic. Furthermore, the Michelson interferometer is not ideal for sub-wavelength displacement measurements as the phase sensitivity is a function of the path length. For example, at the peaks of constructive and destructive interference, the phase sensitivity is zero.

Modern displacement interferometers are based on the heterodyne interferometer by Duke and Gordon from Hewlett-Packard in 1970 [68]. Although similar in principle to a Michelson interferometer, the heterodyne interferometer, overcomes many of the problems associated with the Michelson design. Most importantly, the phase sensitivity remains constant regardless of the path length.

Since the original work in 1970, a wide variety of improvements have been made to the basic heterodyne interferometer, for example, [69]. All of these devices work on the heterodyne principle, where the displacement is proportional to the phase (or frequency) difference between two laser beams. In heterodyne interferometers, the displacement signal is shifted up in frequency which avoids $1/f$ noise and provides immunity from low-frequency light-source intensity variations.

In the original design, the two frequencies were obtained from a He–Ne laser forced to oscillate at two frequencies separated by 2 MHz. However, later designs utilize acousto-optic frequency shifters to achieve a similar result. An example application of a heterodyne interferometer is pictured in Fig. 8.32. Here, the angle and displacement of a linear positioning stage is measured using two interferometers and a single laser source.

A drawback of conventional interferometers is the large physical size and sensitivity to environmental variations which preclude their use in extreme environments such as within a cryostat or high magnetic field. To allow measurement in such environments, the miniature fiber interferometer, pictured in Fig. 8.33a, was developed [70]. The measuring head contains a single-mode optical fiber with a 9 μm core diameter coupled to a collimator lens. Approximately 4 % of the applied light is immediately reflected off the fiber termination and is returned down the fiber, forming the reference beam. The transmitted light passes through the collimator lens and is reflected off the slightly angled target mirror back towards the fiber surface

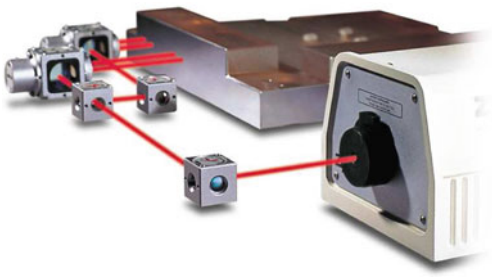


Fig. 8.32 A ZMI™two-axis heterodyne interferometer with a single laser source for measuring the angle and displacement of a positioning stage. Courtesy of Zygo, USA

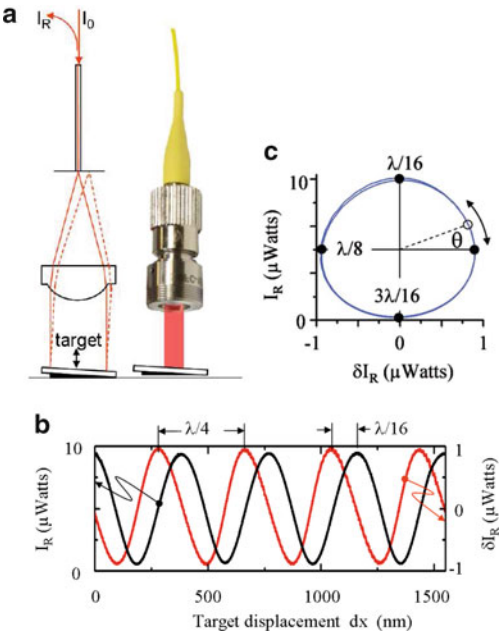


Fig. 8.33 The operating principle of an Attocube FPS miniature fiber interferometer [70], courtesy of Attocube, Germany. In (**a**) the transmitted light is reflected from the mirror, the fiber surface, the mirror again, and is then focused onto the fiber core. The interferogram plotted in (**b**) shows the direct reflected power (*black*) and the quadrature reflected power (*red*) versus displacement. The quadrature signal is obtained by modulating the laser wavelength and demodulating at the receiver. By plotting the power of the direct and quadrature signals (**c**), the direction of travel and sub-wavelength displacement can be resolved

but away from the core. As the fiber surface is a poor reflector, only 4 % of the incident light is reflected from the fiber surface. This reflected light travels back through the lens, is reflected off the mirror and is coupled directly to the fiber core, thus forming a Fabry-Perot interferometer with a cavity length equal to twice the distance between the fiber and mirror.

As the cavity length changes, the two beams interfere so that the reflected power is modulated periodically by the distance as illustrated in Fig. 8.33b. A problem with the basic interferogram is the lack of directional information. To resolve the direction of travel, the light-source wavelength is modulated at a high-frequency and demodulated at the receiver to provide an auxiliary interferogram in quadrature with the original. By considering both the directly reflected power and the demodulated reflected power, the direction of travel and can be deduced from the phase angle shown in Fig. 8.33c.

Since the miniature fiber interferometer is physically separated from the laser and receiver electronics it is both physically small and robust to extreme environments such as high vacuum, cryogenic temperatures, and magnetic fields. Due to the

secondary reflection from the fiber surface, the fiber interferometer is also less sensitive to mirror misalignment compared to some other interferometers.

In general, laser interferometers are the most expensive displacement sensors due to the required optical, laser, and electronic components. However, unlike other sensors, laser interferometers have an essentially unlimited range even though the resolution can exceed 1 nm. Furthermore, the accuracy, stability, and linearity exceed all other sensors. For these reasons, laser interferometers are widely used in applications such as semiconductor wafer steppers and display manufacturing processes. They are also used in some speciality nanopositioning applications that require metrological precision, for example, the metrological AFM described in [15].

Aside from the cost, the main drawback of laser interferometers is the susceptibility of the beam to interference. If the beam is broken, the position is lost and the system has to be restarted from a known reference. The position can also be lost if the velocity of the object exceeds the maximum velocity imposed by the electronics. The maximum velocity is typically a few centimeters per second and is not usually a restriction; however, if the object is subject to shock loads, maximum velocity can become an issue.

The noise of laser interferometers is strongly dependent on the instrument type and operating environment. As an example, the Fabry-Perot interferometer discussed in reference [70] has a $1/f$ noise corner frequency of approximately 10 Hz and a noise density of approximately 2 pm/$\sqrt{\text{Hz}}$. This results in a resolution of approximately 1.6 nm with a 12 kHz bandwidth. Equation (8.29) predicts a resolution of 0.49 nm with a 1 kHz bandwidth. Although the resolution of interferometers is excellent, small range sensors such as capacitive or piezoresistive sensors can provide higher resolution. However, the comparison is hardly fair considering that interferometers have a range in the meters while small range sensors may be restricted to 10 μm or less.

Some manufacturers of interferometers designed for stage metrology and position control include: Agilent, USA; Attocube, Germany (Fibre Interferometer); Keyence, Japan (Fibre Interferometer); Renishaw, UK; Sios, Germany; and Zygo, USA. Instruments from these manufacturers are pictured in Figs. 8.33a and 8.34.

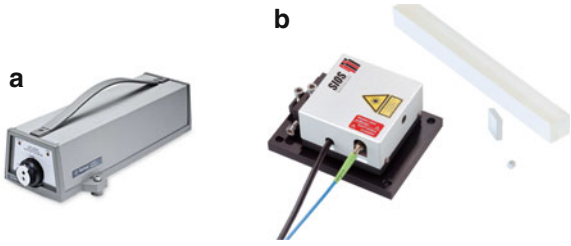


Fig. 8.34 Two commercially available laser interferometers. Photos courtesy of (**a**) Agilent, USA and (**b**) Sios, Germany

8.3.9 *Linear Encoders*

A linear encoder consists of two components, the reference scale and the read-head. The read-head is sensitive to an encoded pattern on the reference scale and produces a signal that is proportional to position. Either the scale or the read-head can be free to move, however, the scale is typically fixed since the read-head is usually lighter.

The earliest form of linear encoder consisted of a bar with a conductive metal pattern, read by a series of metal brushes [21]. Although simple, the constant contact between the brush and scale meant a very limited life and poor reliability.

In the 1950s optical linear encoders became available for machine tools. The reference scales were glass with a photochemically etched pattern. The photolithographic method used to produce the scale resulted in the highest resolution and accuracy at the time.

Although today's optical encoders still produce the highest resolution, other technologies have also become available. Magnetic or inductive linear encoders can not match the absolute accuracy or resolution of an optical scale encoder, however, they are cheaper and more tolerant of dust and contamination. The most common type of encoder is possibly the capacitive encoder found in digital callipers. These devices use a series of conductive lines on the slider and scale to produce a variable capacitor.

The operation of a simple reflective optical encoder is illustrated in Fig. 8.35. Light from a laser diode is selectively reflected from the scale onto a photodetector. As the read-head is moved relative to the scale, the peaks in received power correspond the distance between the reflective bars. In between the peaks, the position can be estimated from the received power. Rather than partial reflection, other gratings contain height profiles that modulate the proximity and thus received power [71].

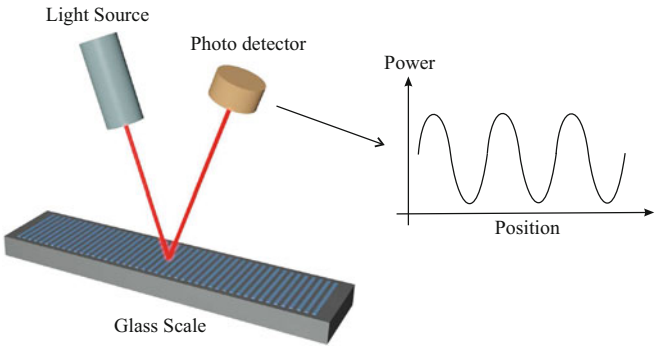


Fig. 8.35 The operation of a simple reflective optical encoder. The peaks in the received power correspond to the distance between reflective bars

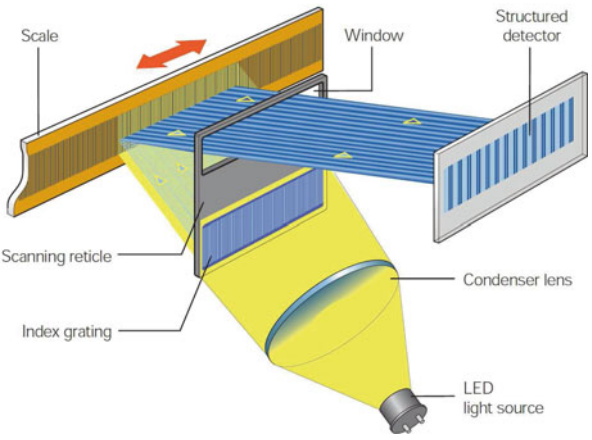


Fig. 8.36 The image scanning technique is used for reference scales with a grating pitch of between 10 and 200 μm. Image courtesy of Heidenhain, Germany

There are two major difficulties with the design illustrated in Fig. 8.35. First, the received power is highly sensitive to any dust or contamination on the scale. Second, it is difficult to determine the direction of motion, particularly at the peaks where the sensitivity approaches zero.

To provide immunity to dust and contamination, commercial optical encoders use a large number of parallel measurements to effectively average out errors. This principle relies on the Moire phenomenon [72] and is illustrated by the image scanning technique shown in Fig. 8.36. In Fig. 8.36 a parallel beam of light is projected onto a reflective scale through a scanning reticle. The reflected Moire pattern is essentially the binary product of the scanning reticle and the scale and is detected by an array of photodetectors. Aside from the immunity to contamination, this technique also provides a quadrature signal that provides directional information.

Optical reference scales are encoded with a geometric pattern that describes either the absolute position or the incremental position. Absolute scales contain additional information that can make them physically larger than incremental scales. Compared to an incremental encoder, an absolute encoder is also typically more sensitive to alignment errors, lower in resolution, slower, and more costly. The benefit of an absolute scale is that the read-head does not need to return to a known reference point after a power failure or read error.

The noise of high resolution optical encoders is described as "jitter" and is typically on the order of 1 nm RMS, or 6 nm peak-to-peak. The overall accuracy is around 5 μm/m [73], however, accuracies as high as 0.5 μm/m are possible with ranges up to 270 mm [74].

The highest resolution optical encoders operate on the principle of interference [74, 75]. The technique involves light that is diffracted through a transparent phase grating in the read-head and reflected from a step grating on the scale [74]. Since this technique operates on the principle of diffraction, extremely small signal periods of down to 128 nm are possible with a resolution on the order of a few nanometers.

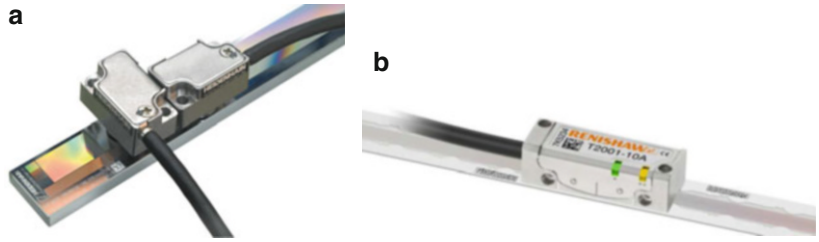


Fig. 8.37 Two commercially available optical linear encoders. Photos courtesy of (**a**) Heidenhain, Germany and (**b**) Renishaw, UK

Table 8.3 Summary of position sensor characteristics

Sensor type	Range	DNR	Resolution (nm)	Max. BW (kHz)	Accuracy
Metal foil	10 – 500 μm	230 ppm	23	1–10	1 % FSR
Piezoresistive	1 – 500 μm	4.9 ppm	0.49	> 100	1 % FSR
Capacitive	10 μm – 10 mm	24 ppm	2.4	100	0.1 % FSR
Electrothermal	10 μm – 1 mm	100 ppm	10	10	1 % FSR
Eddy current	100 μm – 80 mm	10 ppm	1	40	0.1 % FSR
LVDT	0.5 – 500 mm	10 ppm	5	1	0.25 % FSR
Interferometer	m		0.49	> 100	1 ppm FSR
Encoder	m		6	> 100	5 ppm FSR

The dynamic range (DNR) and resolution are approximations based on a full-scale range of 100 μm and a first-order bandwidth of 1 kHz

Other encoder technologies include techniques where the position information is actually encoded into the medium being scanned. Examples of this approach include hard disk drives [76] and MEMS mass storage devices [58].

Companies that produce linear encoders suitable for nanometer scale metrology include: Heidenhain, Germany; MicroE Systems, USA; and Renishaw, UK. Two instruments from these manufacturers are pictured in Fig. 8.37.

8.4 Comparison and Summary

Due to the extreme breadth of position sensor technologies and the wide range of applications, it is extremely difficult to make direct performance comparisons. In many applications, characteristics such as the physical size and cost play a greater role than performance. Nevertheless, it is informative to compare some aspects of performance.

In Table 8.3 the specifications under consideration are the range, the dynamic range, the 6σ-resolution, the maximum bandwidth, and the typical accuracy. Consider the following notes when interpreting the results in Table 8.3:

- The quoted figures are representative of commercially available devices and do not imply any theoretical limits.
- The dynamic range and 6σ-resolution is an approximation based on a full-scale range of 100 μm and a first-order bandwidth of 1 kHz. The low-frequency limit is assumed to be $f_l = 0.01$ Hz.
- The quoted accuracy is the typical static trueness error defined in Sect. 8.2.6.

Metal foil strain gauges are the simplest and lowest cost sensor considered in this study. Due to their size (a few mm^2) strain gauges are suitable for mounting directly on to actuators or stages with a range from 10 to 500 μm. The parameters in Table 8.3 pertain to the example of a two-varying-element bridge discussed in Sect. 8.3.1. Although strain gauges can be calibrated to achieve higher accuracy, it is reasonable to consider an error of 1 % FSR due to drift and the indirect relationship between the measured strain and actual displacement.

Piezoresistive sensors are smaller than metal foil strain gauges and can be bonded to actuators that are only 1 mm long with a range of up to 1 μm. Although the resolution of piezoresistive sensors is very high, the absolute accuracy is limited by nonlinearity, temperature sensitivity, and inexact matching. An error budget of 1 % FSR is typical. Although strain sensors require contact with the actuator or flexural components, they do not introduce forces between the reference and moving platforms, thus, in this sense, they are considered to be non-contact.

Capacitive sensors are relatively simple in construction, provide the highest resolution over short ranges, are insensitive to temperature, and can be calibrated to an accuracy of 0.01 % FSR. However, in general purpose applications where the sensor is not calibrated after installation, alignment errors may limit the accuracy to 1 % FSR. The capacitive sensor parameters under consideration are described in Sect. 8.3.4.

Eddy-current sensors can provide excellent resolution for travel ranges greater than 100 μm. They are more sensitive to temperature than capacitive sensors but are less sensitive to dust and pollutants which is important in industrial environments. The quoted noise and resolution is calculated from the example discussed in Sect. 8.3.6.

LVDT sensors are among the most popular in industrial applications requiring a range from a few millimeter to tens of centimeter. They are simple, have a high intrinsic linearity and can be magnetically shielded. However, they also have a low bandwidth and can load the motion with inertia and friction. The maximum resolution is limited by the physical construction of the transducer which is generally suited to ranges of greater than 1 mm. The bandwidth of LVDT sensors is limited by the need to avoid eddy currents in the core. With an excitation frequency of 10 kHz, the maximum bandwidth is approximately 1 kHz.

Compared to other sensor technologies, laser interferometers provide an unprecedented level of accuracy. Stabilized interferometers can achieve an absolute accuracy exceeding 1 ppm, or in other words, better than 1 um/m. Nonlinearity is also on the order of a few nanometers. Due to the low-noise and extreme range, the dynamic

range of an interferometer can be as high as a few parts per billion, or upwards of 180 dB. The quoted resolution in Table 8.3 is associated with the Fabry-Perot interferometer discussed in Sect. 8.3.8.

Linear encoders are used in similar applications to interferometers where absolute accuracy is the primary concern. Over large ranges, absolute accuracies of up to 5 ppm or 5 μm/m are possible. Even greater accuracies are possible with linear encoders working on the principle of diffraction. The accuracy of these sensors can exceed 1 ppm over ranges of up to 270 mm, which is equivalent to the best laser interferometers.

8.5 Outlook and Future Requirements

One of the foremost challenges of position sensing is to achieve high resolution and accuracy over a large range. For example, semiconductor wafer stages require a repeatability and resolution in the nanometers while operating over a range in the tens of centimeters [13, 14]. Such applications typically use interferometers or high resolution optical encoders which can provide the required performance but can impose a significant cost. Long-range sensors are also becoming necessary in standard nanopositioning applications due to the development of dual-stage actuators [77–80] and stepping mechanisms [81, 82]. Capacitive sensors can be adapted for this purpose by using a periodic array of electrodes [55]. Such techniques can also be applied to magnetic or inductive sensing principles. Due to the increasing availability of long-range nanopositioning mechanisms, an increased focus on the development of cost effective long-range sensors is required.

A need is also emerging for position sensors capable of measuring position at frequencies up to 100 kHz. Applications include: high-speed surface inspection [83, 84]; nanofabrication [3, 4, 12, 85], and imaging of fast biological and physical processes [10, 86–90]. Although, many sensor technologies can provide a bandwidth of 100 kHz, this figure is the 3 dB bandwidth where phase and time delay render the signal essentially useless in a feedback loop. High-speed position sensors are required with a bandwidth in the MHz that can provide accurate measurements at 100 kHz with negligible phase shift or time delay. Due to the operating principle of modulated sensors such as capacitive and inductive sensors, this level of performance is difficult to achieve due to the impractically high carrier frequency requirement. Applications requiring a very high sensor bandwidth typically use an auxiliary sensor for high bandwidth tasks, for example, a piezoelectric sensor can be used for active resonance damping [41, 43]. Technologies such as piezoresistive sensors [91] have also shown promise in high-speed applications since a carrier frequency is not required. Magnetoresistive sensors are also suitable for high-frequency applications if the changes in field strength can be kept small enough to mitigate hysteresis [18, 19].

Due to the lack of cost effective sensors that provide both high-resolution and wide bandwidth, recent research has also considered the collaborative use of

multiple sensors. For example, in reference [40] a piezoelectric strain sensor and capacitive sensor were combined. The feedback loop utilized the capacitive sensor at low frequencies and the piezoelectric sensor at high frequencies. This approach retains the low-frequency accuracy of the capacitive sensor and the wide bandwidth of the piezo sensor while avoiding the drift from the piezo sensor and wide-band noise from the capacitive sensor. The closed-loop noise was reduced from 5 nm with the capacitive sensor to 0.34 nm with both sensors. Piezoelectric force sensors have also been used for high-frequency damping control while a capacitive, inductive, or strain is used for tracking control [31, 41].

Data storage systems are an example application that requires both long range but extreme resolution and increasingly wide bandwidth. In these applications, a media derived position error signal (PES) can provide the requisite range and resolution but not the bandwidth. In reference [58] a MEMs storage device successfully combined the accuracy of a media derived position signal with the speed of an electrothermal sensor. Electrothermal sensors have also been combined with capacitive sensors to reduce the inherent 1/f noise [49]. Multiple sensors can be combined by complementary filters [41] or by an optimal technique in the time-domain [40] or frequency domain [60]. Given the successful applications to date, it seems likely that the trend of multiple sensors will continue, possibly to the point where multiple sensors are packaged and calibrated as a single unit.

References

1. D.Y. Abramovitch, S.B. Andersson, L.Y. Pao, G. Schitter, A tutorial on the mechanisms, dynamics, and control of atomic force microscopes, in *Proceedings of American Control Conference*, pp. 3488–3502, New York City, NY, 2007
2. S.M. Salapaka, M.V. Salapaka, Scanning probe microscopy. IEEE Control Syst. **28**(2), 65–83 (2008)
3. A.A. Tseng (ed.), *Nanofabrication: Fundamentals and Applications* (World Scientific, Singapore, 2008)
4. J.A. Vicary, M.J. Miles, Pushing the boundaries of local oxidation nanolithography: short timescales and high speeds. Ultramicroscopy **108**(10), 1120–1123 (2008)
5. ISO/IEC Guide 98:1993 guide to the expression of uncertainty in measurement, international organization for standardization, 1994
6. ISO 5725 - accuracy (trueness and precision) of measurement methods and results, international organization for standardization, 1994
7. G. Schitter, R.W. Stark, A. Stemmer, Sensors for closed-loop piezo control: strain gauges versus optical sensors. Meas. Sci. Technol. **13**, N47–N48 (2002)
8. S. Devasia, E. Eleftheriou, S.O.R. Moheimani, A survey of control issues in nanopositioning, IEEE Trans. Control Syst. Technol. **15**(5), 802–823 (2007)
9. K.K. Leang, Q. Zou, S. Devasia, Feedforward control of piezoactuators in atomic force microscope systems. IEEE Control Syst. **29**(1), 70–82 (2009)
10. A.J. Fleming, B.J. Kenton, K.K. Leang, Bridging the gap between conventional and video-speed scanning probe microscopes. Ultramicroscopy **110**(9), 1205–1214 (2010). doi:10.1016/j.ultramic.2010.04.016
11. A.J. Fleming, S.S. Aphale, S.O.R. Moheimani, A new method for robust damping and tracking control of scanning probe microscope positioning stages. IEEE Trans. Nanotechnol. **9**(4), 438–448 (2010). doi:10.1109/TNANO.2009.2032418

12. A.A. Tseng, S. Jou, A. Notargiacomo, T.P. Chen, Recent developments in tip-based nanofabrication and its roadmap. J. Nanosci. Nanotechnol. **8**(5), 2167–2186 (2008)
13. H. Butler, Position control in lithographic equipment. IEEE Control Syst. **31**(5), 28–47 (2011)
14. S. Mishra, J. Coaplen, M. Tomizuka, Precision positioning of wafer scanners. segmented iterative learning control for nonrepetitive disturbances. IEEE Control Syst. **27**(4), 20–25 (2007). doi:10.1109/MCS.2007.384130
15. R. Merry, M. Uyanik, R. van de Molengraft, R. Koops, M. van Veghel, M. Steinbuch, Identification, control and hysteresis compensation of a 3 DOF metrological AFM. Asian J. Control **11**(2), 130–143 (2009)
16. Y. Shan, J. Speich, K. Leang, Low-cost IR reflective sensors for submicrolevel position measurement and control. IEEE/ASME Trans. Mechatron. **13**(6), 700–709 (2008). doi:10.1109/TMECH.2008.2005407
17. S. Parkin, X. Jiang, C. Kaiser, A. Panchula, K. Roche, M. Samant, Magnetically engineered spintronic sensors and memory. Proc. IEEE **91**(5), 661–680 (2003). doi:10.1109/JPROC.2003.811807
18. D.R. Sahoo, A. Sebastian, W. Häberle, H. Pozidis, E. Eleftheriou, Scanning probe microscopy based on magnetoresistive sensing. Nanotechnology **22**(14), 145501 (2011)
19. V. Kartik, A. Sebastian, T. Tuma, A. Pantazi, H. Pozidis, D.R. Sahoo, High-bandwidth nanopositioner with magnetoresistance based position sensing. Mechatronics **22**(3), 295–301 (2012)
20. T.R. Hicks, P.D. Atherton, Y. Xu, M. McConnell, *The Nanopositioning Book* (Queensgate Instruments Ltd, Berkshire, UK, 1997)
21. D.S. Nyce, *Linear Position Sensors. Theory and Application* (Wiley, Hoboken, 2004)
22. W.C. van Etten, *Introduction to Noise and Random Processes* (Wiley, West Sussex, England, 2005)
23. R.G. Brown, P.Y.C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering* (Wiley, New York, 1997)
24. A.J. Fleming, A method for measuring the resolution of nanopositioning systems. Rev. Sci. Instrum. **83**(8), 086101 (2012)
25. A.J. Fleming, Measuring picometer nanopositioner resolution, in *Proceedings of Actuator*, Bremen, 2012
26. A.J. Fleming, Estimating the resolution of nanopositioning systems from frequency domain data, in *Proceedings of IEEE International Conference on Robotics and Automation*, St. Paul, MN, pp. 4786–4791, 2012
27. JCGM 200:2008 international vocabulary of metrology - basic and general concepts and associated terms (VIM), 3rd edn. (Bureau International des Poids et Mesures, Paris, 2008)
28. T.-F. Lu, D. Handley, Y.K. Yong, Position control of a 3 DOF compliant micro-motion stage, in *Proceedings of Control, Automation, Robotics and Vision Conference*, vol. 2, pp. 1274–278, 2004 . doi:10.1109/ICARCV.2004.1469029
29. W. Dong, L.N. Sun, Z.J. Du, Design of a precision compliant parallel positioner driven by dual piezoelectric actuators. Sensors Actuators A **135**(1), 250–256 (2007)
30. G. Schitter, P.J. Thurner, P.K. Hansma, Design and input-shaping control of a novel scanner for high-speed atomic force microscopy. Mechatronics **18**(5–6), 282–288 (2008)
31. A.J. Fleming, K.K. Leang, Integrated strain and force feedback for high performance control of piezoelectric actuators. Sensors Actuators A **161**(1–2), 256–265 (2010). doi:10.1016/j.sna.2010.04.008
32. W. Kester, J. Bryant, W. Jung, S. Wurcer, C. Kitchin, Chapter 4. Sensor signal conditioning, in *Op Amp Applications Handbook*, ed. by W. Jung. Analog Devices (Newnes, Oxford, 2002)
33. C.S. Smith, Piezoresistance effect in germanium and silicon. Phys. Rev. **94**(1), 42–49 (1954). doi:10.1103/PhysRev.94.42
34. A. Barlian, W.-T. Park, J. Mallon, A. Rastegar, B. Pruitt, Review: semiconductor piezoresistance for microsystems. Proc. IEEE **97**(3), 513–552 (2009). doi:10.1109/JPROC.2009.2013612

35. C.M. DiBiasio, M.L. Culpepper, Design of a meso-scale six-axis nanopositioner with integrated position sensing, in *Proceedings of Fifth Annual International Symposium on Nanomanufacturing*, Singapore, 2008
36. R. Messenger, Q. Aten, T. McLain, L. Howell, Piezoresistive feedback control of a mems thermal actuator. J. Microelectromech. Syst. **18**(6), 1267–1278 (2009). doi:10.1109/JMEMS.2009.2035370
37. J. Sirohi, I. Chopra, Fundamental understanding of piezoelectric strain sensors. J. Intell. Mater. Syst. Struct. **11**, 246–257 (2000)
38. A.J. Fleming, S.O.R. Moheimani, Control oriented synthesis of high performance piezoelectric shunt impedances for structural vibration control. IEEE Trans. Control Syst. Technol. **13**(1), 98–112 (2005)
39. J. Maess, A.J. Fleming, F. Allgöwer, Simulation of dynamics-coupling in piezoelectric tube scanners by reduced order finite element models. Rev. Sci. Instrum. **79**(1–9), 015105 (2008). doi:10.1063/1.2826428
40. A.J. Fleming, A.G. Wills, S.O.R. Moheimani, Sensor fusion for improved control of piezoelectric tube scanners. IEEE Trans. Control Syst. Technol. **15**(6), 1265–6536 (2008). doi:10.1016/j.mechatronics.2007.07.006
41. A.J. Fleming, Nanopositioning system with force feedback for high-performance tracking and vibration control. IEEE Trans. Mechatron. **15**(3), 433–447 (2010)
42. Y.K. Yong, B. Ahmed, S.O.R. Moheimani, Atomic force microscopy with a 12-electrode piezoelectric tube scanner. Rev. Sci. Instrum. **81**(1–10), 033701 (2010)
43. Y.K. Yong, A.J. Fleming, S.O.R. Moheimani, A novel piezoelectric strain sensor for simultaneous damping and tracking control of a high-speed nanopositioner. IEEE/ASME Trans. Mechatron. (doi: 10.1109/TMECH.2012.2193895, in press). doi:10.1109/TMECH.2012.2193895
44. S.O.R. Moheimani, A.J. Fleming, *Piezoelectric Transducers for Vibration Control and Damping* (Springer, London, 2006)
45. A. Preumont, *Mechatronics, Dynamics of Electromechanical and Piezoelectric Systems* (Springer, Berlin, 2006)
46. L.K. Baxter, *Capacitive Sensors: Design and Applications* (IEEE Press, Piscataway, NJ, 1997)
47. M. Kim, W. Moon, E. Yoon, K.-R. Lee, A new capacitive displacement sensor with high accuracy and long-range. Sensors Actuators A Phys. **130–131**(14), 135–141 (2006). doi:10.1016/j.sna.2005.12.012
48. L.L. Chu, Y.B. Gianchandani, A micromachined 2d positioner with electrothermal actuation and sub-nanometer capacitive sensing. J. Micromech. Microeng. **13**(2), 279–285 (2003)
49. Y.K. Zhu, S.O.R. Moheimani, M.R. Yuce, Simultaneous capacitive and electrothermal position sensing in a micromachined nanopositioner. Electron Device Lett. **32**(8), 1146–1148 (2011)
50. G.T.A. Kovacs, *Micromachined Transducers Sourcebook* (McGraw Hill, Boston, 1998)
51. A.A. Kuijpers, G.J.M. Krijnen, R.J. Wiegerink, T.S.J. Lammerink, M. Elwenspoek, 2d-finite-element simulations for long-range capacitive position sensor. J. Micromech. Microeng. **13**(4), S183–S189 (2003)
52. A.A. Kuijpers, G.J.M. Krijnen, R.J. Wiegerink, T.S.J. Lammerink, M. Elwenspoek, A micromachined capacitive incremental position sensor: part 1. analysis and simulations. J. Micromech. Microeng. **16**(6), S116–S124 (2006)
53. A.A. Kuijpers, G.J.M. Krijnen, R.J. Wiegerink, T.S.J. Lammerink, M. Elwenspoek, A micromachined capacitive incremental position sensor: part 2. experimental assessment. J. Micromech. Microeng. **16**(6), S125–S134 (2006)
54. J.-I. Lee, X. Huang, P. Chu, Nanoprecision MEMS capacitive sensor for linear and rotational positioning. J. Microelectromech. Syst. **18**(3), 660–670 (2009). doi:10.1109/JMEMS.2009.2016275
55. S.-C. Lee, R.D. Peters, Nanoposition sensors with superior linear response to position and unlimited travel ranges. Rev. Sci. Instrum. **80**(4), 045109 (2009). doi:10.1063/1.3123407
56. M.A. Lantz, G.K. Binnig, M. Despont, U. Drechsler, A micromechanical thermal displacement sensor with nanometre resolution. Nanotechnology **16**(8), 1089–1094 (2005)

57. A. Pantazi, A. Sebastian, G. Cherubini, M. Lantz, H. Pozidis, H. Rothuizen, E. Eleftheriou, Control of mems-based scanning-probe data-storage devices, IEEE Trans. Control Syst. Technol. **15**(5), 824–841 (2007). doi:10.1109/TCST.2006.890286
58. A. Sebastian, A. Pantazi, H. Pozidis, E. Elefthriou, Nanopositioning for probe-based data storage. IEEE Control Syst. **28**(4), 26–35 (2008)
59. A. Sebastian, D. Wiesmann, Modeling and experimental identification of silicon microheater dynamics: a systems approach. J. Microelectromech. Syst. **17**(4), 911–920 (2008). doi:10.1109/JMEMS.2008.926980
60. A. Sebastian, A. Pantazi, Nanopositioning with multiple sensors: A case study in data storage. IEEE Trans. Control Syst. Technol. **20**(2), 382–394 (2012). doi:10.1109/TCST.2011.2177982
61. Y. Zhu, A. Bazaei, S.O.R. Moheimani, M. Yuce, Design, modeling and control of a micromachined nanopositioner with integrated electrothermal actuation and sensing, IEEE/ASME J. Microelectromech. Syst. **20**(3), 711–719 (2011)
62. J. Fraden, *Handboook of Modern Sensors: Physics, Designs, and Applications* (Springer, New York, 2004)
63. S. Fericean, R. Droxler, New noncontacting inductive analog proximity and inductive linear displacement sensors for industrial automation. IEEE Sensors J. **7**(11), 1538–1545 (2007)
64. S.D. Roach, Designing and building an eddy current position sensor, Sensors J. Sens. Technol. **15**(9), 56–74 (1998)
65. Q. Li, F. Ding, Novel displacement eddy current sensor with temperature compensation for electrohydraulic valves. Sensors Actuators A Phys. **122**(1), 83–87 (2005). doi:10.1016/j.sna.2005.04.008
66. R. Proksch, J. Cleveland, D. Bocek, Linear variable differential transformers for high precision position measurements, US Patent 7,262,592, 2007
67. P. Hariharan, *Basics of Interferometry*, 2nd edn. (Academic, New York, 2007)
68. J.N. Dukes, G.B. Gordon, A two-hundred-foot yardstick with graduations every microinch. Hewlett-Packard J. **21**(2), 2–8 (1970)
69. G.E. Sommargren, A new laser measurement system for precision metrology, in *Proceedings of Precision Engineering Conference*, Dallas, Texas, 1986
70. K. Karrai, P. Braun, Miniature long-range laser displacement sensor, in *Proceedings of Actuator*, Bremen, Germany, pp. 285–288, 2010
71. A. Khiat, F. Lamarque, C. Prelle, P. Pouille, M. Leester-Schädel, S. Büttgenbach, Two-dimension fiber optic sensor for high-resolution and long-range linear measurements. Sensors Actuators A Phys. **158**(1), 43–50 (2010). doi:10.1016/j.sna.2009.12.029
72. R.S. Sirohi, *Optical Methods of Measurement - Wholefield Techniques* (CRC Press, Taylor and Francis Group, Boca Raton, FL, 2009)
73. FASTRACK high-accuracy linear encoder scale system. data sheet l-9517-9356-01-b (2009). www.renishaw.com
74. Heidenhain exposed linear encoders (2011). www.heidenhain.com
75. J.-Y. Lee, H.-Y. Chen, C.-C. Hsu, C.-C. Wu, Optical heterodyne grating interferometry for displacement measurement with subnanometric resolution. Sensors Actuators A Phys. **137**(1), 185–191 (2007). doi:10.1016/j.sna.2007.02.017
76. B.M. Chen, T.H. Lee, K. Peng, V. Venkatarmanan, *Hard Disk Drive Servo System* (Springer, London, 2006)
77. Y. Michellod, P. Mullhaupt, D. Gillet, Strategy for the control of a dual-stage nano-positioning system with a single metrology, in *Proceedings of Robotics, Automation and Mechatronics*, pp. 1–8, 2006. doi:10.1109/RAMECH.2006.252649
78. L. Chassagne, M. Wakim, S. Xu, S. Topçu, P. Ruaux, P. Juncar, Y. Alayli, A 2d nanopositioning system with sub-nanometric repeatability over the millimetre displacement range. Meas. Sci. Technol. **18**(11), 3267–3272 (2007)
79. A.J. Fleming, Dual-stage vertical feedback for high speed-scanning probe microscopy. IEEE Trans. Control Syst. Technol. **19**(1), 156–165 (2011). doi:10.1109/TCST.2010.2040282
80. J. Zheng, A. Salton, M. Fu, Design and control of a rotary dual-stage actuator positioning system. Mechatronics **21**(6), 1003–1012 (2011). doi:10.1016/j.mechatronics.2011.03.011

81. C.-L. Chu, S.-H. Fan, A novel long-travel piezoelectric-driven linear nanopositioning stage, Precis. Eng. **30**(1), 85–95 (2006). doi:10.1016/j.precisioneng.2005.05.002
82. R. Merry, M. Maassen, M. van de Molengraft, N. van de Wouw, M. Steinbuch, Modeling and waveform optimization of a nano-motion piezo stage. IEEE/ASME Trans. Mechatron. **16**(4), 615–626 (2011). doi:10.1109/TMECH.2010.2050209
83. G. Borionetti, A. Bazzalia, R. Orizio, Atomic force microscopy: a powerful tool for surface defect and morphology inspection in semiconductor industry. Eur. Phys. J. Appl. Phys. **27**(1–3), 101–106 (2004)
84. A. Humphris, M. McConnell, D. Catto, A high-speed atomic force microscope capable of video-rate imaging. Microsc. Anal. SPM Suppl. **20**, 29–31 (2006)
85. A. Ferreira, C. Mavroidis, Virtual reality and haptics for nanorobotics. IEEE Robot. Autom. Mag. **13**(3), 78–92 (2006)
86. G.E. Fantner, G. Schitter, J.H. Kindt, T. Ivanov, K. Ivanova, R. Patel, N. Holten-Andersen, J. Adams, P.J. Thurner, I.W. Rangelow, P.K. Hansma, Components for high speed atomic force microscopy. Ultramicroscopy **106**(2–3), 881–887 (2006)
87. M. Kobayashi, K. Sumitomo, K. Torimitsu, Real-time imaging of DNA-streptavidin complex formation in solution using a high-speed atomic force microscope. Ultramicroscopy **107**(2–3), 184–190 (2007)
88. G. Schitter, K.J. Åström, B.E. DeMartini, P.J. Thurner, K.L. Turner, P.K. Hansma, Design and modeling of a high-speed AFM-scanner. IEEE Trans. Control Syst. Technol. **15**(5), 906–915 (2007)
89. L.M. Picco, L. Bozec, A. Ulcinas, D.J. Engledew, M. Antognozzi, M. Horton, M.J. Miles, Breaking the speed limit with atomic force microscopy. Nanotechnology **18**(4), 044030(1–4) (2007)
90. T. Ando, T. Uchihashi, T. Fukuma, High-speed atomic force microscopy for nano-visualization of dynamic biomolecular processes. Prog. Surf. Sci. **83**(7–9), 337–437 (2008)
91. E. Guliyev, T. Michels, B. Volland, T. Ivanov, M. Hofer, I. Rangelow, High speed quasi-monolithic silicon/piezostack spm scanning stage. Microelectron. Eng. **98**(0), 520–523 (2012). doi:10.1016/j.mee.2012.07.059

Chapter 9
MEMS Nanopositioners

Jason J. Gorman

9.1 Introduction

Nanopositioning mechanisms, or nanopositioners, have evolved quickly over the last few decades due to a growing need for nanoscale precision motion control in applications including microscopy, lithography, nanomanufacturing, and optics. The critical differentiator between nanopositioners and other motion stages is that nanopositioners are capable of nanoscale positioning resolution, which is generally interpreted as a few nanometers and below. Motion stages with mechanical elements that have friction, such as screws, linear bearing slides, and rotational bearings, are not capable of nanoscale resolution. The backlash caused by friction limits the precision in these mechanisms to tens of nanometers, and complex control systems are often required to reach this level. As a result, nanopositioner designs typically use a flexure mechanism to guide the motion of the stage, which exhibit no friction since the motion is based on structural bending, thereby providing continuous smooth movements.

The simplest nanopositioner is composed of a flexure mechanism and an actuator that drives the mechanism. In most high-precision applications, nanopositioners also include sensors for measuring the motion, which are used for feedback control and other error compensation methods. Among the many nanopositioner designs in the literature, the combination of piezoelectric actuators, a metal flexure mechanism, and capacitive or optical sensors is by far the most common. This combination can be seen in numerous linear, multi-axis, and rotational nanopositioners [1–3] and has been the basis of the majority of commercially available nanopositioners [4–6].

J.J. Gorman (✉)
National Institute of Standards and Technology, 100 Bureau Drive,
Stop 8212, Gaithersburg, MD 20899-8212, USA
e-mail: gorman@nist.gov

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_9

These motion stages have radically improved the performance of many precision instruments, including scanning probe microscopes [7] and optically lithography tools for semiconductor electronics manufacturing [8].

In tandem with the evolution of nanopositioners over the last few decades, microelectromechanical systems (MEMS) have emerged to create a major industry and have enabled numerous technology advances including automotive airbag deployment, inertial navigation for unmanned vehicles, and smartphones. Most MEMS are composed of combinations of actuators, flexure mechanisms, and sensors, as seen in accelerometers, gyroscopes, and resonators. Furthermore, these devices have been shown to be capable of exceptional precision compared to their macroscale counterparts. As a result, MEMS are an obvious technology platform for extending the capabilities of nanopositioners. This opportunity was realized early on in the development of both nanopositioners and MEMS by Akamine et al. [9] who developed a microscale piezoelectric cantilever for scanning tunneling microscopy that can move with three degrees of freedom. Since then, MEMS nanopositioners have progressed rapidly due to exciting applications that are not possible with macroscale nanopositioners, including multi-tip scanning probe microscopy (SPM) for high-bandwidth imaging, lithography, and data storage.

This chapter will provide a review of the state of the art for MEMS nanopositioners. The next section discusses the motivation for using MEMS in nanopositioning, including the components within a nanopositioner, the performance requirements for these systems, the advantages in using MEMS to create high-performance motion stages, and the applications that are driving the development of this technology. This is followed by a review of the actuators and sensors that have been demonstrated in MEMS nanopositioners, with an emphasis on the advantages and disadvantages related to precision motion control. An assessment of the performance of MEMS nanopositioners to date in terms of resolution, range, bandwidth, and size is then presented. Finally, the chapter concludes with thoughts on the current limitations and future directions for MEMS nanopositioners.

9.2 Motivation for MEMS Nanopositioners

Nanopositioners are composed of actuators, a flexure mechanism, sensors, and a control system. The block diagram in Fig. 9.1 describes the interactions between these components. For a single-axis system, a control voltage, u, is amplified and applied to the actuator. This causes the actuator to exert a force, F, on the flexure mechanism, thereby causing a displacement, x. The displacement is then measured with the sensor and the sensing signal is processed by the readout electronics to produce a voltage, V_x, which is proportional to the displacement. The sensor voltage is then sent to the control system, which generates the control voltage, u, based on the selected algorithm. Open-loop control is also possible but it typically yields significantly reduced precision.

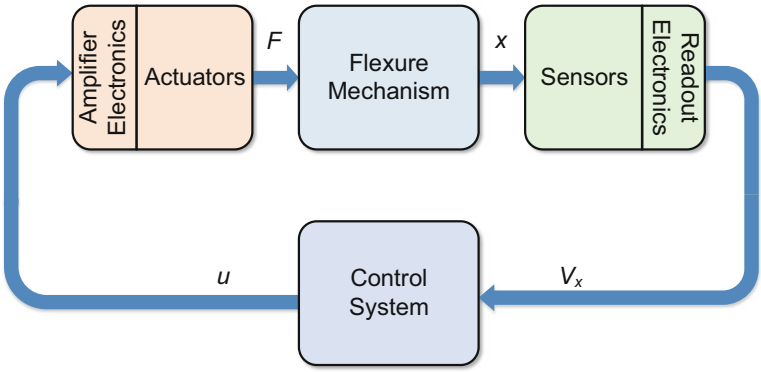


Fig. 9.1 Anatomy of a nanopositioner. Block diagram showing the components within a nanopositioner and the interactions between the components. u = control voltages, F = forces applied by actuators, x = displacements of MEMS nanopositioner, V_x = measured displacements in voltage

Actuators for controlled motion and sensors for displacement measurement are widely used in MEMS including those found in accelerometers, gyroscopes, microphones, and pressure sensors. As a result, MEMS technologies provide a compelling approach to implementing the device anatomy shown in Fig. 9.1. The main question is whether MEMS nanopositioners can match the performance obtained by macroscale nanopositioners. While there is considerable variation in the performance attained with macroscale nanopositioners, the motion resolution can be as good as 0.05 nm, the range can be as high as a few hundred micrometers, and the dynamic range is typically above 10^5. The achievable motion bandwidth ranges from tens of hertz to several kilohertz and is inversely related to the motion range. There are hard constraints between these four metrics (resolution, range, dynamic range, and bandwidth) but they can be optimized for a given application. For example, scanning tunneling microscopy (STM) usually requires greater resolution than atomic force microscopy (AFM). Therefore, a nanopositioner for STM may have motion resolution on the order of tens of picometers but with a range of only a few micrometers, whereas a nanopositioner for AFM may have sub-nanometer resolution and a range around 100 μm.

The motivation for the development of MEMS nanopositioners is a combination of advantages related to performance, functionality, and new applications that are only accessible with small mechanisms. Scaling down in size presents one set of advantages. First, MEMS have high resonant frequencies compared to larger systems, resulting in greater motion bandwidth. Second, MEMS can be integrated on a silicon chip with other technologies, such as electronics and photonics, thereby providing a path to low-power embedded systems with high functionality. Finally, arrays of nanopositioners can be fabricated on a single chip and controlled independently. The main disadvantages in scaling down in size are the resulting payload size and mass limits.

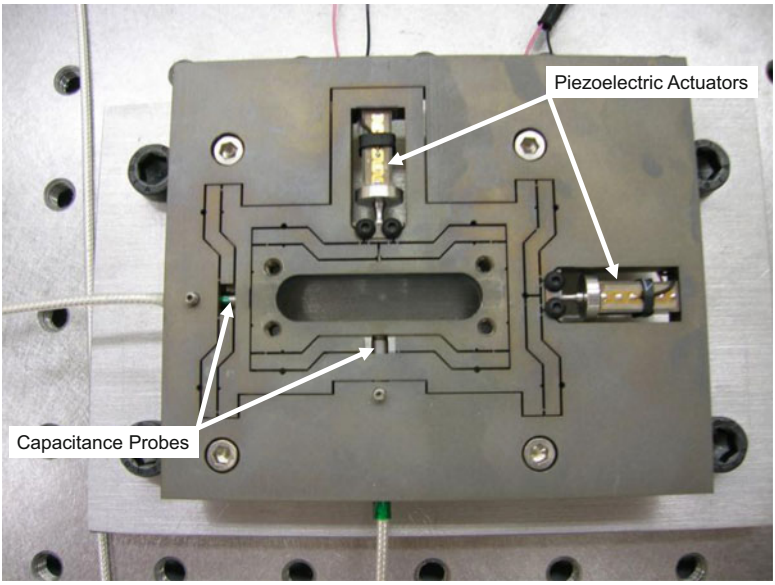


Fig. 9.2 An *XY* nanopositioner with piezoelectric actuators, a nested flexure mechanism with motion amplification, and capacitance probes (J.J. Gorman, S. Bergna, N.G. Dagalakis, unpublished)

Another set of advantages results from the integration and economy of scale inherent in MEMS fabrication processes. A representative macroscale *XY* nanopositioner is shown in Fig. 9.2. It includes two piezoelectric stack actuators, a nested *XY* flexure mechanism that amplifies the actuator motion, and two capacitive position sensors. The flexure mechanism is fabricated using wire electrodischarge machining (wire EDM), which is a slow, expensive, and serial process. The actuators and sensors are assembled into the mechanism by hand, adding additional time and cost to fabrication. Additionally, misalignment between components is inevitable, causing off-axis forces and erroneous position measurements, thereby reducing the accuracy of motion. In comparison, MEMS fabrication uses parallel processes, the mechanisms are typically monolithic, and assembly is not required. Therefore, the cost is significantly reduced and motion errors due to misalignment of sensors and actuators are all but eliminated.

Figure 9.3 shows an *XY* MEMS nanopositioner with electrothermal actuators and a nested flexure mechanism that amplifies motion, similar to the mechanism shown in Fig. 9.2. Comparing Figs. 9.2 and 9.3, the advantages of MEMS fabrication are evident. The MEMS nanopositioner is fabricated in a single layer of single-crystal silicon. Complex flexure mechanisms, such as the multi-lever arm design with circular notch flexure joints can be fabricated with ease using deep reactive ion etching (DRIE). The actuators and flexure mechanism are directly integrated and the only misalignments possible between the two are caused by lithography errors,

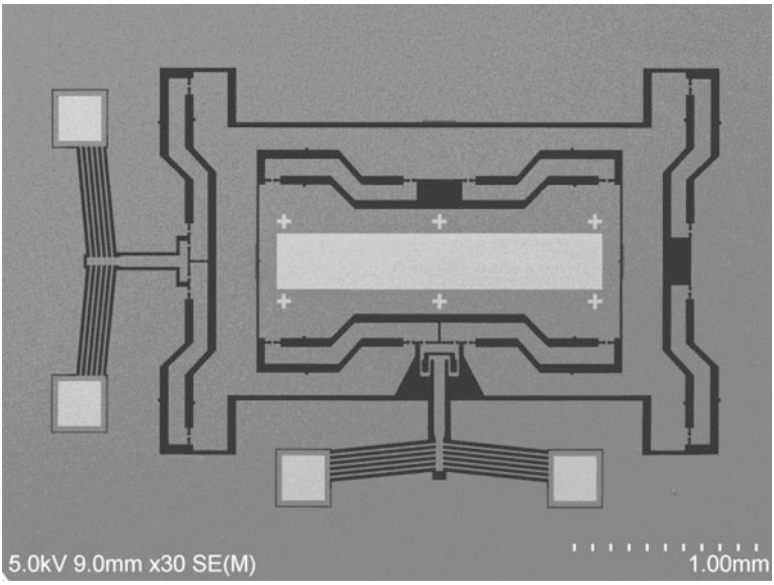


Fig. 9.3 An *XY* MEMS nanopositioner with electrothermal actuators and a nested flexure mechanism with motion amplification, similar to that shown in Fig. 9.2 (J.J. Gorman, S. Bergna, N.G. Dagalakis, unpublished)

which are typically negligible. This is in stark contrast to the complex couplings necessary to assemble the actuators and sensors in the macroscale nanopositioner in Fig. 9.2.

The materials available in MEMS fabrication also provide significant advantages over those common in macroscale nanopositioners. Single-crystal silicon is the most widely used material for MEMS nanopositioners due to several factors. It is broadly available with widely varying specifications, its material properties are well known, and there are many well-established etch processes available, making it a relatively easy material to use in MEMS fabrication. Silicon has a much higher Young's modulus and a lower coefficient of thermal expansion than the materials used in macroscale nanopositioners (i.e., metals), resulting in stiffer structures and improved thermal stable. Monolithic integration in silicon further improves thermal stability by avoiding mismatched material properties as found in assembled nanopositioners.

There are several applications driving the development of MEMS nanopositioners with SPM being the most prominent and well explored. By definition, SPM [10], including STM and AFM, requires a nanopositioner to scan the probe or sample with exceptional precision. There is a continual push to increase the scan rate in SPM in order to reduce drift in surface images, maximize data collection, and observe dynamic nanoscale phenomena. As described above, MEMS are capable of high motion bandwidth making them an obvious solution to the SPM scanning problem. Akamine et al. [9] were the first to demonstrate a MEMS multi-axis

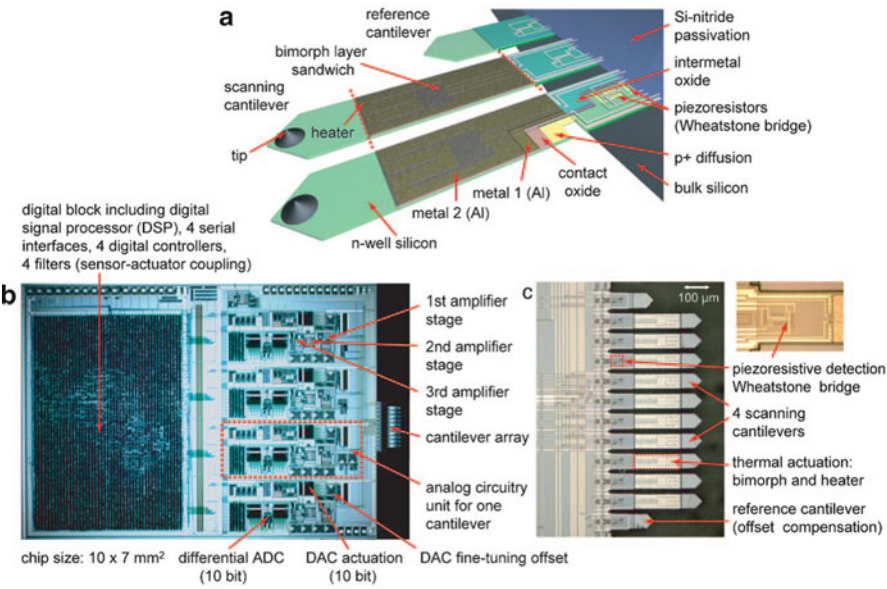


Fig. 9.4 Active AFM cantilever array. (**a**) Schematic of cantilever array showing actuation and sensing elements, (**b**) embedded readout electronics, and (**c**) fabricated cantilever array (S. Hafizovic et al. © PNAS [18])

scanner for STM, achieving atomic resolution with a piezoelectric cantilever. This work was followed by the development of multi-axis electrostatic nanopositioners with integrated tunneling tips that were designed for multi-tip STM measurements [11–13].

Most of the research on MEMS scanners has focused on AFM due to the broader applications compared to STM and the simpler imaging mechanism (i.e., tip force vs. tunneling current). Due to their small size and high level of device integration, MEMS have provided a path for active AFM cantilever arrays with independent force sensing and *Z* axis actuation [14–18] (see Fig. 9.4). In addition to the increased motion bandwidth due to scaling down in size, cantilever arrays provide multiplexed measurements that increase the image size and speed of image collection by increasing the number of cantilevers. While multiple tips can be controlled in the vertical direction (e.g., 10 cantilevers [14]), scanning of the cantilever array along the *X* and *Y* axes is performed with a macroscale nanopositioner. Arrays of independent AFM scanners (*XYZ*) would be capable of significantly higher throughput due to their higher motion bandwidth. As a step in this direction, an *XY* MEMS nanopositioner has been used to image metal patterns on the scanner with resolution and bandwidth that is competitive with traditional AFM [19–22]. Most recently, a fully integrated on-chip AFM with three axes of motion and embedded force sensing has been shown to be capable of imaging nanoscale step heights with

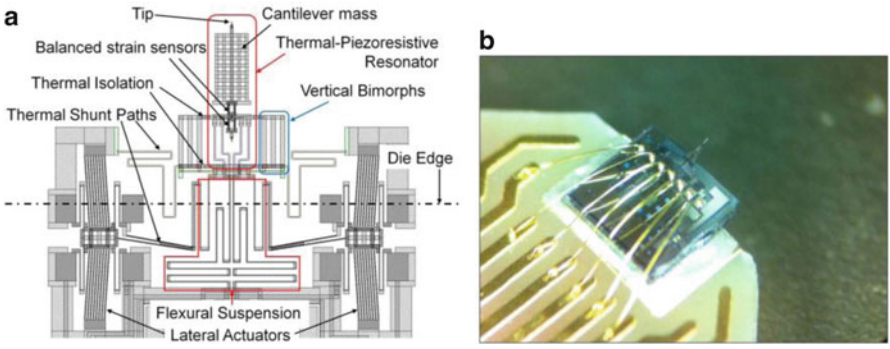


Fig. 9.5 On-chip AFM. (**a**) Schematic showing three motion axes with electrothermal actuation and cantilever with piezoresistive strain sensing and embedded tip, (**b**) fabricated on-chip AFM (N. Sarkar et al. © IEEE [24])

3 nm pk-pk vertical resolution [23, 24] (see on-chip AFM in Fig. 9.5). Applications for on-chip AFM beyond nanoscale imaging include probe-based data storage [25] and nanolithography [14].

Another active application area for MEMS nanopositioners is the manipulation of nanostructures for nanomanufacturing and material testing. A number of nanomanipulator concepts have been explored in which probes with nanoscale sharpness are used to push and tweeze nanostructures [26–29]. Nanopositioners have also been used to measure the mechanical properties of nanowires and carbon nanotubes by attaching them to the motion stage and pulling on them while observing the event in a scanning electron microscope (SEM) or transmission electron microscope (TEM) [30–32]. By scaling down the material testing experiment, the accuracy of the results can be improved due to a significant reduction in the distance between sensors and actuators and because they can more readily fit in small sample chambers as found in TEMs. Many other applications for MEMS nanopositioners are expected in the areas of adaptive optics (e.g., tunable lens arrays), photonics (e.g., laser cavity stabilization), and physics research (e.g., optical and atomic traps), among others.

Returning to Fig. 9.1, it is clear that for MEMS nanopositioners to be successful in the applications above, each of the components (actuators, flexure mechanism, sensors, controller) must perform at a level comparable to their macroscale counterparts. There has been considerable research on actuators and sensors for MEMS nanopositioners, which will be discussed in the following sections. Control system design for MEMS nanopositioners varies greatly depending on the actuator, flexure mechanism, and sensor selections, and the application of interest. As a result, this subject is discussed in detail in the following chapter. The final component in a MEMS nanopositioner, the flexure mechanism, is critical to its performance and requires careful design to achieve stable motion over the desired range. However, flexure mechanism design at the microscale is largely the same as at the macroscale (e.g., see [33, 34]), so this topic is not covered here. However, it is worth noting the variety of mechanism designs that have been used to date.

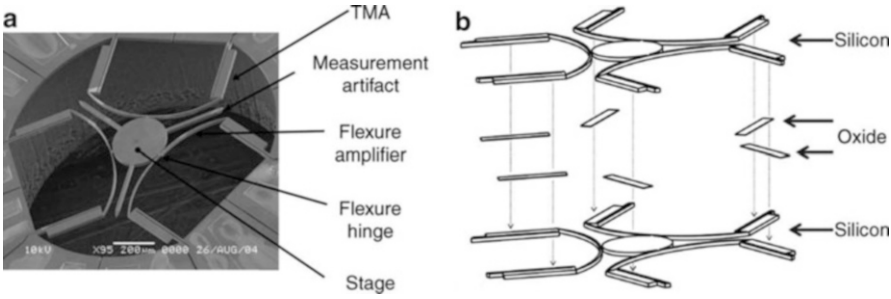


Fig. 9.6 The μHexFlex nanopositioner. This mechanism uses electrothermal actuators and a two-level flexure mechanism to generate six-axis motion. (**a**) Fabricated mechanism, (**b**) exploded diagram showing two levels of thermal actuators. (S.-C. Chen and M.L. Culpepper © Elsevier [47])

Early mechanisms relied heavily on cantilevered beams [9, 14–18] and clamped beams [35]. Double parallelogram flexures were then found to constrain the motion better than simple beam structures, resulting in pure translational motion along the degree of freedom of interest with minimal off-axis motion [35–38]. Multi-axis MEMS nanopositioners require more complex mechanisms to properly constrain the motion. A number of different designs have been used for *XY* stages, including beam structures [12, 19, 39, 40], a four-bar flexure [41], and nested structures with circular notch flexure hinges [28, 42] (see Fig. 9.3). Other multi-axis mechanisms of interest include *XYZ* stages [43–45] (see Fig. 9.8), an $XY\theta$ stage [46], and even a six-axis parallel mechanism [47], as shown in Fig. 9.6. Almost all of these flexure mechanism designs work equally well at the microscale as at the macroscale. One of the biggest challenges in developing an adequate flexure mechanism for MEMS is finding a design that is compatible with the fabrication process. In many cases, this limits the mechanism to two-dimensional geometry in the plane of the substrate, which can make it difficult to properly constrain both in-plane and out-of-plane motions simultaneously. Other important design challenges include the trade-off between range and bandwidth (i.e., fundamental resonant frequency), cross-talk between motion axes, and achieving linearity over the entire workspace.

9.3 Actuation

Actuators for nanopositioning must meet several stringent performance requirements. Smooth continuous motion with position fluctuations on the order of 1 nm or below is required. Sufficient force must be generated to move the flexure mechanism over a range from 1 to 100 μm, depending on the application. In general, a motion bandwidth of at least several hundred hertz is necessary but significantly higher bandwidth is typically expected since this is a driving motivation for using MEMS nanopositioners. Depending on the level of integration required for the application, this must all be done within a volume from 0.1 to around 5 mm^3. Actuation in

the plane of the supporting substrate and out of the plane are both required for many applications of MEMS nanopositioners, such as SPM, since they depend on multi-axis motion. There are numerous books [48, 49] and review articles [50–52] that describe the design and modeling of MEMS actuators and sensors in detail. Therefore, the goal of this section and the following section on sensors is not to provide an in-depth description on how various actuators and sensors operate. Rather, the focus is on highlighting the approaches that have been used in MEMS nanopositioners and discussing their advantages and disadvantages within this context. Among the many intriguing approaches for actuation in MEMS, almost all of the MEMS nanopositioners to date rely on electrostatic, electrothermal, or piezoelectric actuation, which are discussed below. The most obvious omission from this discussion is electromagnetic actuation, which can be very effective at the macroscale, particularly linear Lorentz force actuators. However, due to poor scaling and difficulties in fabricating microscale rare-earth magnets, electromagnetic actuation has received little attention with respect to MEMS nanopositioners, with a few exceptions (e.g., see [53] and its reference list).

9.3.1 Electrostatic Actuators

Electrostatic actuation is the most common approach for MEMS nanopositioners, as well as for MEMS in general. The design and modeling of electrostatic actuators are well understood, making them easy to integrate into mechanisms and achieve a predictable level of performance. An electrostatic actuator is a capacitor in which one side of the capacitor can move because it is attached to a flexure mechanism. When a voltage is applied across the capacitor a charge is generated, resulting in an electrostatic force that is proportional to the voltage squared. The three most common types of electrostatic actuators are the comb actuator, the parallel-plate actuator, and the parallel-plate actuator with comb structure, as shown in Fig. 9.7.

The electrostatic comb actuator (Fig. 9.7a) uses a pair of interdigitated combs to generate a force in the direction parallel to the comb fingers (i.e., the x-axis). Numerous MEMS nanopositioners have used this actuator due to its simplicity of design, well-characterized force model, and considerable motion range [12, 19, 35–41, 43, 44, 46]. An example of an *XYZ* MEMS nanopositioner with comb actuators for the X and Y axes is shown in Fig. 9.8. The force can be optimized by minimizing the gap between the fingers, g, and maximizing the thickness thickness, t. Fabrication on silicon-on-insulator wafers using DRIE is the best method for achieving these high aspect ratio structures. Although the force is proportional to the voltage squared, the actuator response can easily be linearized by canceling the nonlinearity with a square root function, either with an analog circuit or digital signal processor. The motion range of a comb actuator can be several hundred micrometers through careful design of the connected flexure mechanism [54, 55], which is exceptionally large for MEMS nanopositioners. Most comb actuators are designed for in-plane motion, but out-of-plane motion (z axis) can also

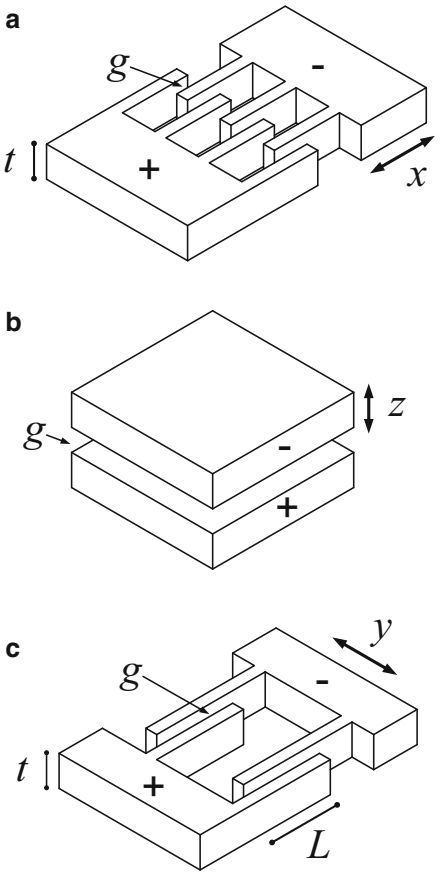


Fig. 9.7 Common electrostatic actuator designs: (**a**) Comb actuator, (**b**) parallel-plate actuator, and (**c**) parallel-plate actuator with a comb structure

be achieved by using the levitation effect [13] or an offset in the vertical position of the two combs [56, 57]. This results in a much smaller motion range, typically on the order of a few micrometers.

The electrostatic parallel-plate actuator (Fig. 9.7b) has been used at many scales and its application in MEMS predates the use of comb actuators. Several actuator geometries for in-plane [11] and out-of-plane motion [43, 58–61] have been used in MEMS nanopositioners. An example is shown in Fig. 9.8, where the parallel-plate actuator drives the Z axis of an XYZ MEMS nanopositioner. The design of this actuator is simple and easy to fabricate. However, the force generated by a parallel-plate actuator is nonlinearly related to the displacement of the moving plate, resulting in a pull-in instability, which is generally not an issue for comb actuators. The instability causes the moving plate to snap in to the stationary plate when the displacement due to an applied voltage exceeds one third of the nominal gap, g. The motion range before pull-in can be improved through mechanical design (e.g., see [62]), but it is still small compared to comb actuators. When g is made larger to accommodate a greater range, the required input voltage necessary to achieve a long

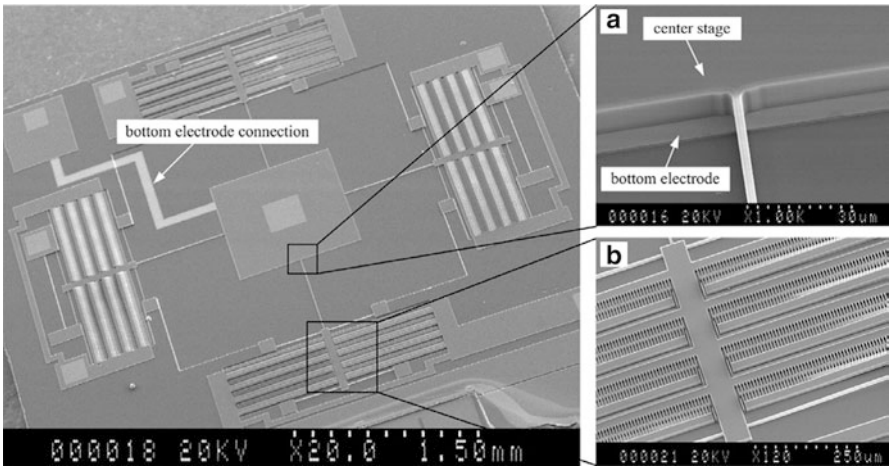


Fig. 9.8 An *XYZ* MEMS nanopositioner that uses electrostatic comb actuators for the *X* and *Y* axes and an electrostatic parallel-plate actuator for the *Z* axis (X. Liu, K. Kim, and Y Sun © IOP [43])

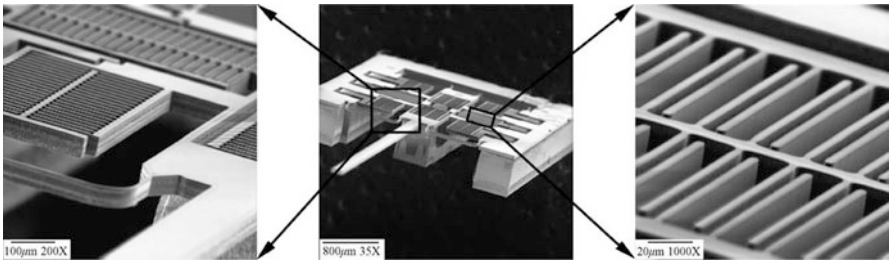


Fig. 9.9 An *XY* MEMS nanopositioner that uses electrostatic parallel-plate actuators with a comb structure (Y. Sun et al. © Elsevier [64])

range increases and becomes impractical, thereby setting the motion limit to only a few micrometers. The parallel-plate approach is typically used for out-of-plane actuation but it can be implemented for in-plane motion by using a comb structure with many parallel-plate actuators working in parallel, as shown in Figs. 9.7c and 9.9 [63, 64].

Among the many positive attributes for electrostatic actuation, possibly the most attractive is that it has a flat dynamic response (i.e., no resonances). The dynamics of the nanopositioner are defined solely by the flexure mechanism, making it possible to achieve the desired motion bandwidth through straightforward modeling of the electrostatic actuator and the structural dynamics of the flexure mechanism. The greatest challenge in using an electrostatic actuator is that the ratio of the actuation force over the in-plane actuator area is low in comparison to other actuators. As a result, large actuators are necessary to generate an adequate motion range in most cases. This limits the electrostatic actuation to applications in which the deeply scaled integration of MEMS nanopositioners arrays is not required.

9.3.2 *Electrothermal Actuators*

Electrothermal actuators are the second most popular type of actuator for MEMS nanopositioners. In these actuators, a current is passed through an electromechanical structure with an appropriate resistance (typically between 100 Ω and 1 kΩ). This results in Joule heating causing the structure to expand due to the increase in temperature. If designed correctly, the thermal expansion results in linear motion that is proportional to the voltage squared since the temperature increase is proportional to the power dissipation. By properly constraining the electromechanical structure, the actuator motion can be significantly larger than that achieved through linear expansion alone. Thermal expansion provides continuous smooth motion that can easily be controlled through the actuation voltage.

Almost all electrothermal actuators can be defined as being one of following two types: (1) Joule heating through the bulk of a single-material actuator or (2) Joule heating of stacked, thermally mismatched materials (i.e., bimorph actuation). In the first type, current flows through the bulk of the structure and the actuator uniformly expands and contracts across its cross-section. The two most common geometries for this type are the chevron actuator and the U-shape actuator, as shown in Fig. 9.10. The chevron actuator [65] has a symmetric geometry with angled beams connected to a center shuttle that transmits the force. When current flows through the beams, they heat up and their length increases, causing the center shuttle to move. The resulting output motion is related to the drive voltage, the geometry of the beams,

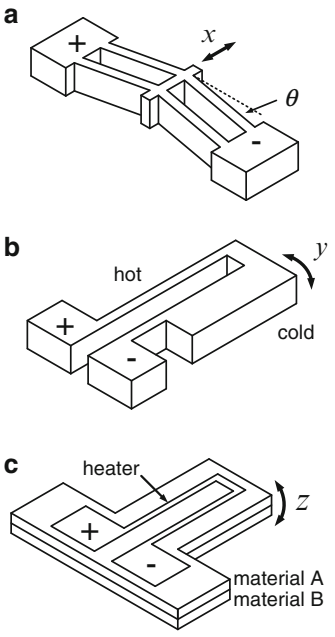


Fig. 9.10 Common electrothermal actuator designs. (**a**) chevron actuator, (**b**) U-shape actuator, and (**c**) thermal bimorph actuator

and the beam angle. A smaller beam angle provides larger motion for the same drive voltage but also reduces the actuator stiffness, thereby reducing the force that can be transmitted. Numerous MEMS nanopositioners have relied on this actuator [65–72] due to its simplicity of design and because is generates highly linear motion.

The U-shape actuator [73] is composed of a narrow beam (hot arm) and a wide beam (cold arm) that are connected in a continuous loop. When a current flows through the loop, the hot arm heats up and expands, while the cold arm experiences a much smaller increase in temperature due to its lower resistance. This mismatch in expansion causes motion at the tip of the actuator along an arc. Nanopositioning typically requires linear motion so the arc motion of the U-shape actuator is unacceptable in most cases. This has limited it use in comparison to the chevron actuator. Linear motion can be achieved though by coupling two actuators together as shown in [47] (see Fig. 9.6). The geometry shown in Fig. 9.10b and in [47, 73] is for in-plane motion. However, the principle of using a hot and cold arm has been demonstrated for out-of-plane motion as well [74].

Bimorph actuators that use Joule heating are composed of two or more materials that have different coefficients of thermal expansion and a heater, which is a conductive structure that is driven electrically (see Figs. 9.10c and 9.11). When a current is applied to the heater, the actuator heats up and moves due to internal stress caused by the mismatch in thermal expansion and the thermal gradient. This approach is most commonly used with cantilevers [16–18, 75, 76] because large motion can be generated at the cantilever tip due to small strain at the base from thermal expansion. Furthermore, it is relatively easy to stack the materials necessary to achieve a large mismatch in their expansion. As with the U-shape actuator,

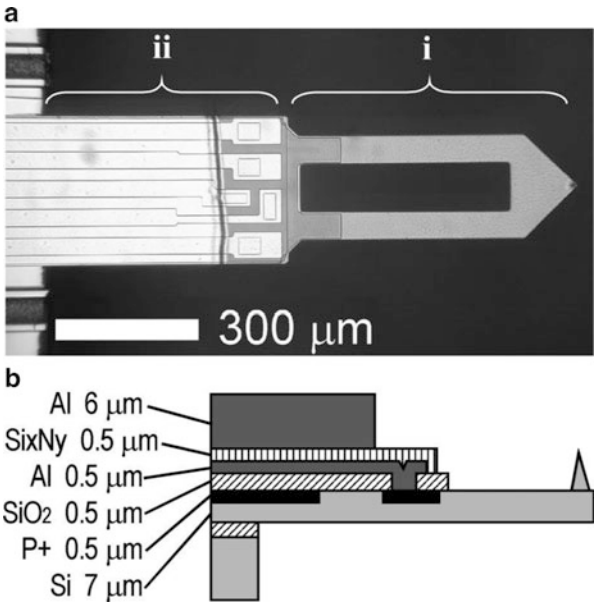


Fig. 9.11 Thermal bimorph cantilever. (**a**) Section *ii* is the thermal bimorph actuator and section *i* contains a piezoresistive sensor, (**b**) schematic of cantilever cross-section (T. Akiyama et al. © AIP [75])

bimorph cantilevers generate motion along an arc making them less attractive for nanopositioning. However, the thermal bimorph concept has been implemented to generate in-plane linear motion by using a CMOS (complementary metal-oxide semiconductor) fabrication process to create actuators composed of dielectric structures and metal heaters [23, 24, 77]. These in-plane actuators use beam bending resulting from bimorph actuation but the heaters are strategically placed within the beams and multiple bimorph beams are coupled together. This design cancels out the rotational motion typically found in bimorph actuators, thereby providing linear output motion. As a result, this design can be used in applications similar to that of the chevron actuator.

While there is considerable variation in the motion range of thermal actuators, the maximum displacement is around 20 μm. The reduced range compared to electrostatic actuators is frequently compensated using a flexure mechanism with motion amplification [28, 42]. Similar to electrostatic actuators, the displacement of electrothermal actuators is proportional to the square of the applied voltage. However, when this input nonlinearity is canceled using an analog circuit or digital signal processor, their dynamic response is nearly linear with the exception of temperature-dependent material properties. The chevron and U-shape actuators are straightforward to fabricate since they only require a single device layer, making these devices attractive for silicon-on-insulator processes.

In addition to ease of fabrication, electrothermal actuators are robust mechanisms that can withstand a large number of heating cycles, are generally unaffected by surface contamination, and are less prone to electrical breakdown compared to electrostatic actuators. The biggest disadvantage of electrothermal actuators is their bandwidth, which is limited by their thermal time response. Typically, a bandwidth in the range of 100 Hz can be achieved, although this is dependent on size, heat sinking, and whether the actuator is operating in air or vacuum. The bandwidth can be extended to some degree through open-loop control [75, 78]. Even with open-loop control, the bandwidth is limited by the maximum drive voltage that can be sustained without damage and the rate at which the actuator can lose heat. This drawback of electrothermal actuation must be strongly considered for a given application.

9.3.3 Piezoelectric Actuators

Multi-layer piezoelectric actuators, such as those shown in Fig. 9.2, are by far the most commonly used actuators in macroscale nanopositioners. Interestingly, piezoelectric actuators have found very limited use in MEMS nanopositioners. This is largely due to the small piezoelectric coefficients that can be achieved (in this case, the d_{33} coefficient), thereby requiring stacked actuators with thick layers. For example, typical multi-layer piezoelectric actuators made with lead zirconate

titanate (PZT) can achieve approximately 1 μm of displacement at 120 V for a 1 mm thick material layer. Clearly, this approach does not scale well since a 1 μm thick layer would result in only 1 nm of displacement. As a result, linear actuators in which the motion has a one-to-one correspondence to the expansion of a piezoelectric material are generally not found in MEMS with the exception of resonators, which can operate with exceptionally small motion.

One approach for achieving adequate motion range with a MEMS piezoelectric actuator is to use a bimorph cantilever design [9, 14, 15]. Expansion of the piezoelectric material causes the cantilever to bend, resulting in vertical motion of the cantilever tip. Displacements of tens of micrometers can be achieved with this approach. However, linear motion cannot be generated with this configuration, limiting its application. An intriguing approach to solving this problem uses multiple bimorph actuators to generate in-plane linear motion through out-of-plane bending, resulting in a range of a few micrometers [79]. In addition to the challenge of generating linear motion, the deposition of high-quality thin-film piezoelectric materials with properties comparable to those found in macroscale materials remains an open area of research. As a result, piezoelectric actuators are not currently practical for MEMS nanopositioners but this may change in the near future.

9.4 Sensing

The most commonly used sensors in macroscale nanopositioners are capacitive sensors and strain gauge sensors, including both metal and piezoresistive types. Laser interferometry is also used when exceptional accuracy and stability are required and cost and size are not a concern. These displacement sensors must meet a set of stringent performance requirements similar to those described for actuators in the previous section: resolution of 1 nm or below, range from 1 to 100 μm, and bandwidth of a few kilohertz. In addition, linearity and thermal stability are critical since they directly affect the motion accuracy. However, it is more difficult to define these metrics and their values are largely unreported for MEMS in the literature.

Among the MEMS sensors for linear displacement measurement, either capacitive, thermal, or piezoresistive sensors have been used in almost all MEMS nanopositioners to date, as described below. The most intriguing option that has yet to receive adequate attention is optical sensing, including interferometric [80] and diffractive optics [81] sensors. With the continued integration of photonics into MEMS, optical sensors in MEMS nanopositioners are expected over the next decade.

9.4.1 Capacitive Sensors

Capacitive sensing has been integral to a number of MEMS sensor technologies including accelerometers, gyroscopes, pressure sensors, and force sensors, and their mechanical and electronic design is well developed. This is because they can achieve exceptional displacement resolution and are easy to fabricate in standard MEMS processes. As a result, capacitive sensors are the most common displacement sensors used in MEMS nanopositioners. Capacitive sensors convert a displacement into a change in capacitance, which is measured using readout electronics. The variable capacitor designs that are typically used are the same as the electrostatic actuators shown in Fig. 9.7: the comb sensor [29, 66, 82–84], parallel-plate sensor [58], and parallel-plate sensor with comb structure [63, 64]. Similar to electrostatic actuators, the comb sensor has a linear relationship between capacitance and displacement, whereas parallel-plate designs are nonlinear. Since sensor linearity is critical in nanopositioning, the comb sensor is highly preferred. Sub-nanometer resolution has been achieved using both single-ended comb sensors, with [29, 66] and without [83] motion amplification, and with differential comb sensors [84]. Comb sensors have many of the same advantages as comb actuators, including large range, flat dynamic response, high bandwidth, and ease of fabrication.

The readout electronics that measure the change in capacitance when the nanopositioner moves are just as important as the structure of the sensor. The most sensitive readout electronics for capacitive sensors are AC bridge and switched capacitor circuits [85]. These circuits can achieve a noise floor better than 0.1 aF/rt-Hz, making it possible to reach sub-nanometer resolution, and a dynamic range above 10^5, which are ideal characteristics for nanopositioning. Furthermore, these circuits are compatible with CMOS electronics [85, 86], making it possible to integrate the MEMS and electronics on a single chip or co-integrate in a single package.

As with electrostatic actuators, the main disadvantage of comb sensors is the area required to attain a sufficiently sensitive device. An important recent innovation for capacitive sensing in nanopositioners is the use of the same comb structure for both actuation and sensing [87, 88]. This reduces the area required for actuation and sensing by half, thereby increasing its bandwidth by reducing the mass of the nanopositioner. In one configuration, the actuation signal and the AC signal used to measure the comb capacitance are separated in the frequency domain and filtering is used to lock in on the sensor signal [87]. The frequency of the AC signal is selected to be well above the mechanical resonances of the structure so that it does not impart a measurable motion. Alternatively, the comb has been used as a tunable capacitor within a resonant LC circuit [88]. The resonant frequency of the circuit is modulated by the motion of the nanopositioner, which can be calibrated to attain a sensitivity in Hz/nm. The change in frequency of the resonator can be measured using a phased locked loop, providing a high bandwidth sensor output. Actuation and sensing on the same comb has yet to be compared to the performance of independent actuation and sensing channels and warrants further investigation.

9.4.2 Thermal Sensors

Over the last decade, thermal sensors have emerged as a viable alternative to capacitive sensors for displacement measurement in MEMS [71, 72, 83, 89–92]. In its simplest form, a thermal sensor uses heat conduction between an electrothermal heater and a moving heat sink, typically the motion stage, as a measure of position. The heater is most commonly made of doped silicon and is shaped to evenly distribute the heat over the area of the heater that interacts with the motion stage. The heater and the motion stage interact through a small gap (<5 μm) that conducts heat away from the heater. As the overlapping area between the motion stage and heater changes due to displacement, the temperature of the heater changes, resulting in a change in the resistance of the heater. The change in resistance is measured with readout electronics and the output voltage is calibrated using an accurate measurement of the motion of the nanopositioner. The sensor and motion stage can be stacked vertically [72, 89, 91] or they can interact in the plane of motion [71, 83, 90]. In most cases, the thermal sensor and actuator moving the motion stage are separated to avoid coupling of their dynamics. However, it is possible to intentionally couple an electrothermal actuator to a thermal sensor, which significantly reduces the area consumed by the mechanism [91].

Thermal sensors are typically measured in pairs in which the overlapping area increases for one heater and decreases for the other for a given motion of the stage. Assuming a constant voltage across the heaters, the current flowing through one heater increases and the other decreases. By measuring the difference between the two signals, sensor drift due to environmental temperature fluctuations and sensor aging can be canceled out. This differential configuration of thermal sensors for measuring the motion of a moving stage is shown in Fig. 9.12. Trans-impedance amplifiers are used to convert the sensor currents to voltage, which are then subtracted from each other to attain the drift-free sensor output. Interestingly, when directly compared to capacitive sensing, thermal sensors yielded a lower noise floor above 4 Hz in one implementation [83], demonstrating the applicability of this approach to nanopositioning. Thermal sensors, similar to the piezoresistive sensors discussed in the next subsection, have high $1/f$ noise compared to capacitive sensors due to fluctuations in the conductivity of the heaters. This issue has been addressed through AC modulation of the heater voltage and measuring the output voltage with a lock-in amplifier at the modulation frequency, which has been shown to improve the noise floor at low frequency by a factor of 2 [92].

Among the many promising features of thermal sensors, they have been shown to be capable of a resolution of 0.5 nm, a motion range of tens of micrometers, and 10 kHz bandwidth [89]. Additionally, the majority of the sensors to date are no bigger than 400 $\mu m \times 200$ μm, which is considerably smaller than capacitive sensors. The fabrication of most existing thermal sensor designs is straightforward, particularly the in-plane design shown in Fig. 9.12 since it uses a standard silicon-on-insulator process. Sensor linearity is design dependent, with some demonstrating high linearity [89] and others with strong nonlinearity [71]. The largest drawback of

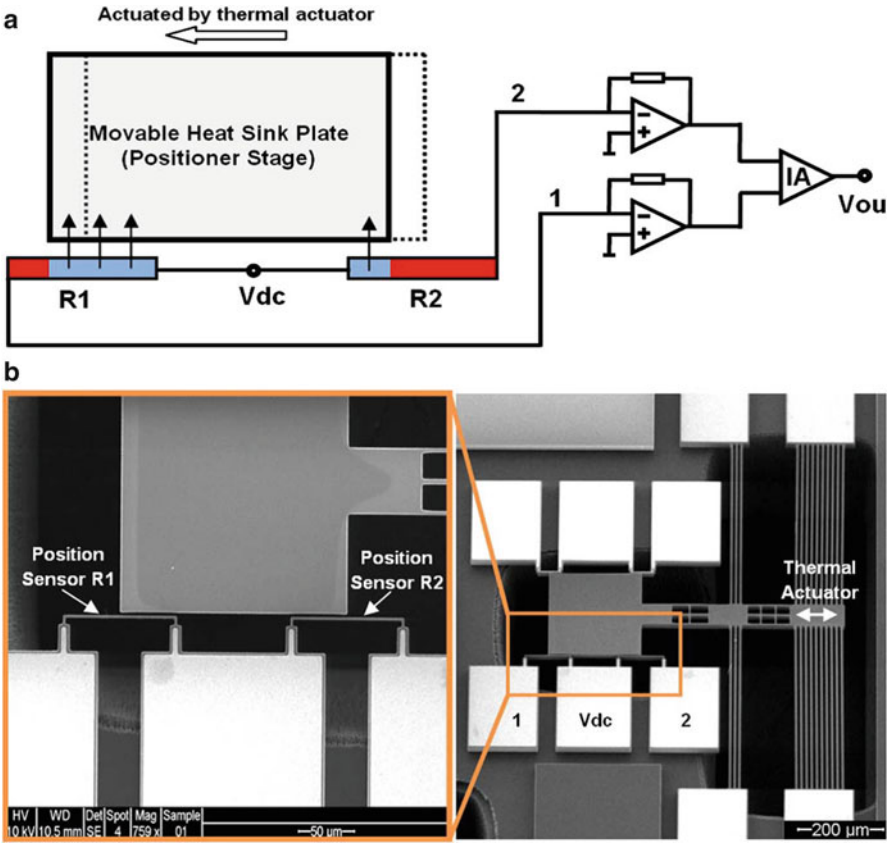


Fig. 9.12 Thermal displacement sensor with differential readout electronics. (**a**) Schematic of approach for differential readout of two thermal sensors, (**b**) scanning electron micrographs of the thermal sensors, motion stage, and the electrothermal actuator that moves the stage (Zhu et al. © IEEE [90])

thermal sensors is that they act as heat sources that can cause significant asymmetric thermal expansion in the mechanism. MEMS nanopositioners require excellent motion accuracy and stability and thermal expansion could degrade performance in these areas. This issue requires further study to determine if it will impact MEMS nanopositioner applications.

9.4.3 Piezoresistive Sensors

Piezoresistive sensing was one of the first sensing mechanisms to be used in MEMS and has been integral to many of the most common sensor products including

pressure sensors, accelerometers, gyroscopes, and microphones [93]. Piezoresistors are similar to metal strain gauges in that their resistance changes due stress. However, metal strain gauges change resistance only due to geometric changes caused by the stress. Piezoresistors, which are fabricated from semiconducting materials, most commonly silicon, also change resistance due to a stress-induced change in the bandgap of the material, which is typically a much larger effect than the change due to geometric effects. Piezoresistive sensors are frequently used in low-precision macroscale nanopositioners because they are compact, allowing them to be directly bonded to piezoelectric stack actuators. Unfortunately, their resolution is typically one to two orders of magnitude worse than capacitive sensors. The first MEMS nanopositioners to use piezoresistive sensors were cantilever arrays designed for SPM [15–18]. One example is shown in Fig. 9.4, in which four piezoresistors are located at the base of each AFM cantilever within the array. The piezoresistors are embedded in the cantilever through selective doping of silicon. This approach measures the induced strain due to the out-of-plane motion of the cantilever tip rather than measuring the tip motion directly.

Measurement of in-plane motion using piezoresistive sensors has been achieved by using suspended silicon structures that extend or compress when the nanopositioner moves. For example, long slender beams were used as piezoresistors to measure the displacement of a motion stage over a 12 μm range [94]. The piezoresistors were fabricated in the same silicon layer as the nanopositioner, as shown in Fig. 9.13, greatly simplifying fabrication since actuation and sensing are formed in the same etch step and selective doping is not required. Similar to the simultaneous actuation and sensing approach used with electrostatic comb actuators [87, 88], piezoresistive sensing can be combined with electrothermal actuation [70, 95]. The beams in a chevron electrothermal actuator experience a change in resistance during actuation due to stress and a change in resistivity due to the change in temperature. As a result, the change in resistance can be measured during actuation to attain the actuator displacement. However, due to coupling between actuation and sensing, the response is nonlinear and it is not straightforward to separate these coupling effects.

Another approach for integrating piezoresistors into flexural structures uses a CMOS–MEMS fabrication process that is also compatible with making independent electrothermal actuators [23, 24], as described in the previous section. In this case, the piezoresistors are composed of polysilicon and are embedded in a silicon dioxide structural layer. Both in-plane and out-of-plane motion can be measured by strategically placing the piezoresistors to measure the strain related to the desired degree of freedom.

There are only a handful of cases where piezoresistive sensing has been applied successfully to MEMS nanopositioners so it is difficult to quantify the expected resolution and range. Hafizovic et al. [18] demonstrated a 1 nm resolution over a 1 μm range, while Bazaei et al. [94] achieved 5 nm resolution over a range of 12 μm. While the resolution is promising, the dynamic range reported to date is quite low, with the best result among the included references being 8000. The resolution is generally limited by the high level of $1/f$ noise and Johnson noise that is universal

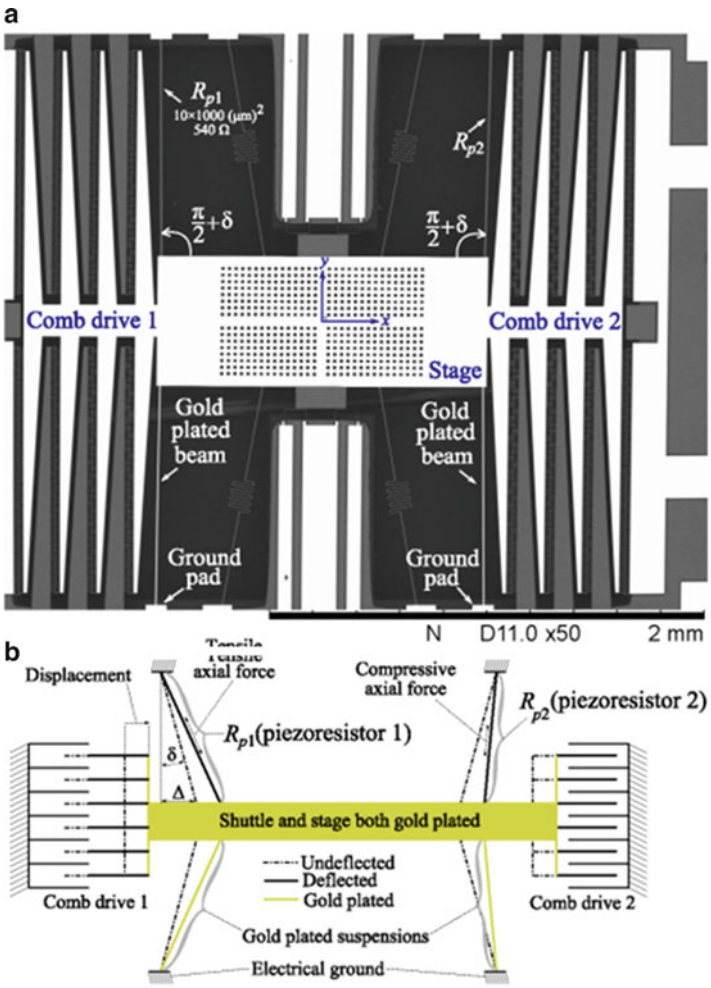


Fig. 9.13 MEMS nanopositioner with piezoresistive sensors. (**a**) Scanning electron micrograph showing thin silicon piezoresistors (R_{p1} and R_{p2}) connected to the motion stage, (**b**) Schematic describing the strain in each piezoresistor due to stage motion (Bazaei et al. © IEEE [94])

to all resistors. The use of AC readout electronics can eliminate the contribution from the $1/f$ noise, but the Johnson noise is still a major contributor to the noise floor.

The biggest advantages of piezoresistive sensors are their small size, thereby minimizing the overall area consumed by a mechanism, and their high bandwidth, which is typically limited by the structural dynamics of the flexure mechanism and the readout electronics. Attaining piezoresistors with a high gauge factor (i.e., high sensitivity) typically requires selective ion implantation or diffusion doping, complicating the fabrication process. However, adequate displacement resolution

has been achieved without selective doping [94], so a high gauge factor may not be required in some applications. Thermal drift is another well-known problem with piezoresistors that could be a perceived issue for MEMS nanopositioners. However, this problem is easily mitigated by using a full Wheatstone bridge to readout the resistance change and by including reference piezoresistors in the design that do not experience stress. In general, piezoresistive sensors are an intriguing approach for displacement measurement in MEMS nanopositioners that requires more research to determine whether sub-nanometer resolution can be achieved over tens of micrometers of motion.

9.5 Performance Assessment

A number of different MEMS nanopositioners have been described in the previous sections. An open question is whether these mechanisms can perform as well as macroscale nanopositioners. Although this is a difficult question to answer since many of the performance characteristics of interest are rarely reported, this section provides an assessment of the performance of existing MEMS nanopositioners and compares the results with the typical performance achieved at the macroscale. The focus of this comparison is on the aggregate rather than highlighting the top performing mechanisms and does not examine the effect of actuator and sensor selection on performance. The following subsection discusses the methodology for the assessment and is followed by the presentation and discussion of the results.

9.5.1 Methodology

Four performance metrics have been selected for this assessment: range, resolution, bandwidth, and in-plane area. Data was cataloged from the relevant references previously cited in this chapter and [96–103], with 59 references providing usable data points. A minimum of the in-plane area and one other metric were required for a reference to be included in this analysis. The range is defined as either a stated value or the maximum displacement presented in the reference. Resolution is defined as either the stated RMS resolution or is calculated from the displacement noise floor in $m/Hz^{1/2}$ using the bandwidth. Due to $1/f$ noise and other noise sources, there can be a considerable difference between the RMS noise and that calculated using the noise floor and bandwidth. However, this is unavoidable for this assessment since resolution is so rarely stated explicitly. The bandwidth is taken as the lowest value among the stated bandwidth, the half-power point of the amplitude frequency response, and the first resonant frequency. Finally, the in-plane area is typically extracted from provided images and is defined as the width multiplied by the length. For mechanisms with multiple degrees of freedom, the width and length are divided by the degrees of freedom to provide a fair comparison with single axis devices.

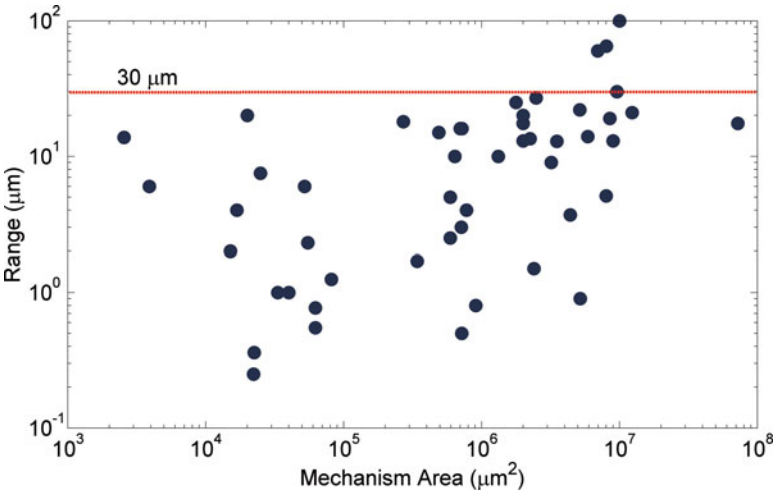


Fig. 9.14 Range as a function of in-plane mechanism area for MEMS nanopositioners

9.5.2 Results

Among the four metrics of interest, range is the most widely reported. Comparing motion range versus in-plane area, as shown in Fig. 9.14, there is a wide distribution in both of these metrics. Most MEMS nanopositioners have a range below 30 μm with some exceptions going to 100 μm and beyond (some outliers were not included in the figures). This is shorter than the average range for macroscale nanopositioners but is expected for MEMS and 30 μm is sufficient for many of the applications of interest. One clear trend is that larger area is required for the longest-range mechanisms, which is obvious since more force and larger flexure mechanisms are required for large range. More interesting is that below 10 μm the same range is achieved by mechanisms ranging in area from 10^4 to 10^7 μm^2. This is due in part to the trade-offs in actuator selection and the common use of flexure mechanisms that are not optimized for the expected range, resulting in excessively large dimensions.

Resolution is the single most important performance metric since it determines whether a motion stage can truly achieve nanoscale motion precision. Unfortunately, achieving a clear assessment of resolution is challenging since multiple definitions are used including the RMS value of the displacement noise and the displacement noise floor. Furthermore, it is commonly not addressed in publications; less than half of the references providing range data also include resolution. Finally, independent confirmations of resolution, as done with a laser interferometer at the macroscale, are rare. Even with these challenges, there are some useful insights attained by looking at resolution as a function of in-plane area, as shown in Fig. 9.15. Similar to the range data, there is a wide distribution for resolution, spanning more than two orders of magnitude. Most of the available data is for larger mechanisms (≈1 mm^2 or larger) but there is no obvious trend indicating a relationship between scale and

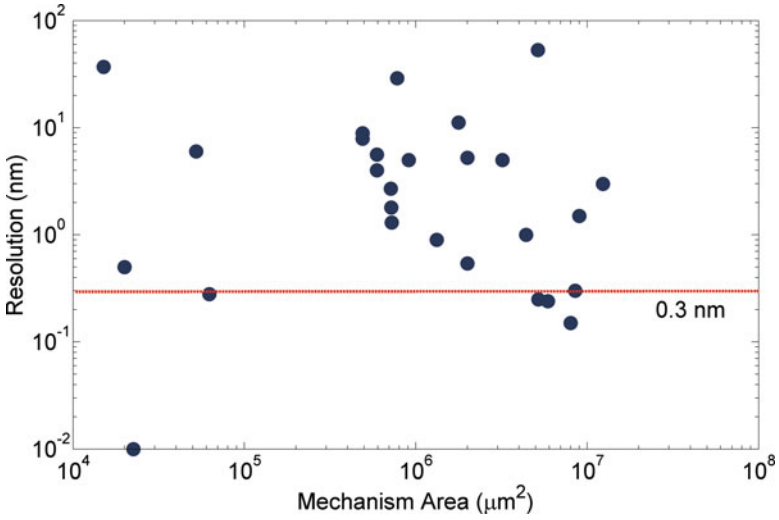


Fig. 9.15 Resolution as a function of in-plane mechanism area for MEMS nanopositioners

resolution. In terms of magnitude, there are a handful of papers reporting sub-nanometer resolution but the majority of mechanisms are above 2 nm, making it difficult to define them as nanopositioners.

As previously discussed, the dynamic range (DR) (i.e., range/resolution) for macroscale nanopositioners is typically greater than 10^5. Assuming that the range is below 30 μm as described in the previous paragraph and a DR = 10^5, the resolution should be less than 0.3 nm as indicated by the dashed line in Fig. 9.15. It is clear that only a few reported designs can perform at this level. Taking another look at this issue, range versus resolution is plotted in Fig. 9.16 along with contour lines for dynamic range. The majority of reported designs perform below DR = 10^4 and no MEMS nanopositioner considered in this analysis has achieved a DR $\geq 10^5$. This is the single greatest limitation in the state of the art. The resolution needs to be improved to meet that achieved at the macroscale while also increasing the dynamic range beyond 10^5 (i.e., improve resolution while not decreasing range).

As expected, bandwidth is the metric in which MEMS nanopositioners show exceptional performance, as shown in Fig. 9.17. Most reported mechanisms have bandwidth greater than 100 Hz and more than half have demonstrated bandwidth greater than 1 kHz. This bandwidth range is sufficient for many applications and can clearly be further optimized considering some examples above 10 kHz have been demonstrated. The data in Fig. 9.17 shows that, as expected, bandwidth is linked to in-plane area with the greatest bandwidth occurring on the low end of the spectrum (200 kHz at 6×10^4 μm^2) and the smallest bandwidth at the high end (20 Hz at 7×10^6 μm^2). Based on Figs. 9.14, 9.15, and 9.17, the majority of MEMS nanopositioners to date have an area greater than 2.5×10^5 μm^2 (i.e., equivalent to a 500 μm square). Reduction in this area would improve bandwidth and facilitate large-scale integration of arrays for scanning probe applications.

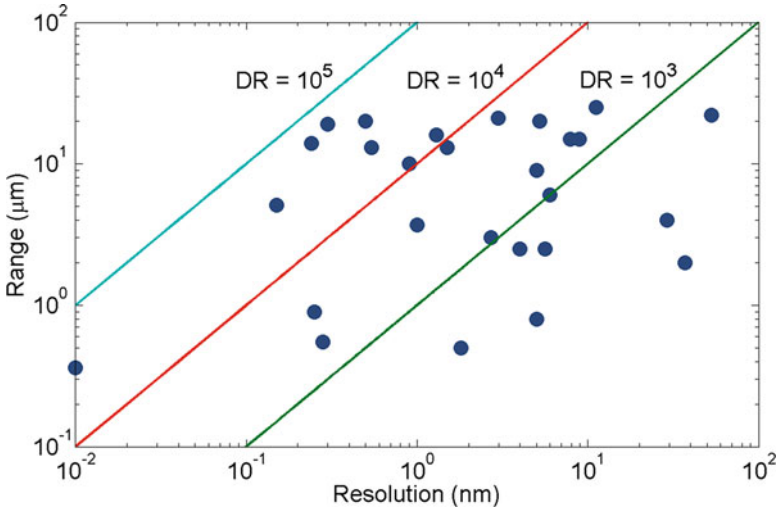


Fig. 9.16 Range as a function of resolution for MEMS nanopositioners. *DR* dynamic range

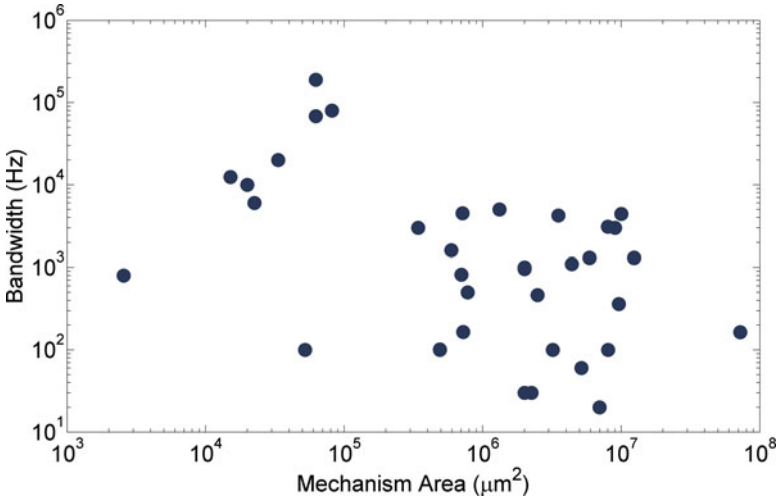


Fig. 9.17 Bandwidth as a function of in-plane mechanism area for MEMS nanopositioners

While the presented data and analysis is clearly not complete it does provide some important insights into the state of the art. The range and bandwidth of existing mechanisms are adequate for known applications and in some cases the bandwidth is exceptional compared to non-MEMS approaches. Unfortunately, the resolution and dynamic range found in this analysis are both worse than that found in macroscale nanopositioners by an order of magnitude or more in most cases. This provides a clear direction for future research with a goal of achieving resolution <0.1 nm RMS and DR > 10^5.

9.6 Conclusion

MEMS nanopositioners have emerged over the last two decades as a highly capable class of mechanisms that merges precision engineering design principles with MEMS sensors and actuators. They are already being used successfully in SPM and probe-based data storage, with products expected within the next few years, and hold promise in many other applications (e.g., nanomanufacturing, nanomaterial testing, biomedical, photonics). This chapter has provided an overview of the critical components within these mechanisms, provided discussion on the advantages and disadvantages of the most common actuators and sensors, and presented a meta-analysis of their performance to date.

Among the possible actuator technologies of interest, electrostatic, electrothermal, and piezoelectric actuators were identified as the most compatible and widely used with MEMS nanopositioners. Piezoelectric actuators still require substantial research to reach the current performance levels achieved by the other two approaches, particularly with respect to piezoelectric material deposition and actuator design. Selection between electrostatic and electrothermal actuation can be determined almost solely on bandwidth requirements. Bandwidth above a few hundred hertz requires electrostatic actuation in most cases because the thermal response of electrothermal actuators is typically too slow. Otherwise, electrothermal actuators are preferred over electrostatic actuators when their low bandwidth can be tolerated because they require lower actuation voltage, consume less area, and are more reliable.

The most commonly used sensing approaches for MEMS nanopositioners are capacitive, thermal, and piezoresistive sensors. While capacitive sensing has been most commonly used in precision MEMS applications, thermal and piezoresistive sensing have been shown to achieve resolution and range performance that is approaching that of capacitive sensing in some situations. Furthermore, unlike the actuators discussed above, all three approaches provide adequate bandwidth for almost all nanopositioning applications. As a result, sensor selection must be based on available in-plane area (capacitive sensing takes the most area), the intended fabrication process (piezoresistive sensing is maximized by local ion implantation or diffusion doping), and the operating environment (thermal sensors are less sensitive in vacuum), among other factors. A detailed study of the three sensor types that compares their performance limits under equivalent constraints and operating conditions would be highly beneficial to future nanopositioner designs.

The presented performance assessment of existing MEMS nanopositioners revealed two deficiencies in the state of the art. First, the achievable resolution in the majority of devices reported in the study is worse by an order of magnitude or more in comparison to macroscale nanopositioners. A resolution of 0.1 nm RMS is commonly achieved in macroscale nanopositioners but most MEMS nanopositioners are above 1 nm. Second, the dynamic range among the studied mechanisms was found to be an order of magnitude worse than their macroscale counterparts ($>10^5$) in most cases. Somewhat less critical, it was found that that the majority of MEMS

nanopositioners have an in-plane area greater than 2.5×10^5 μm^2, or a 500 μm square. This size limitation should be improved because it currently restricts the capabilities of nanopositioner arrays due to poor fill factor. The other performance metrics of interest, range, and bandwidth have been shown to be adequate in many of the existing mechanisms.

Looking to the future, research should be focused on improving resolution and dynamic range since these metrics are the best indicators of usability in nanopositioning applications and the state of the art is inadequate. If 0.1 nm RMS resolution and a dynamic range above 10^5 can be achieved, MEMS nanopositioners will be capable of replacing macroscale nanopositioners in scanning probe applications. Another area that warrants more attention is the integration of scanner arrays with independent *XYZ* motion. Previous work on parallel operation has focused on cantilever arrays with a single degree of freedom in each cantilever. Recent developments in on-chip AFMs [24] have opened the door for independent scanning probe arrays that will be significantly faster and more fault tolerant than existing cantilever arrays. Embedded integration of the mechanisms and electronics will be critical to the success of such systems. Other important areas of research include analysis of positioning stability, cross talk between axes, and the limitations of actuation and readout electronics. With sustained efforts in these research areas, it is expected that MEMS nanopositioners will outperform macroscale nanopositioners within the following decade.

References

1. J.W. Ryu, D.-G. Gweon, K.S. Moon, Optimal design of a flexure hinge based XYφ wafer stage. Precis. Eng. **21**, 18–28 (1997)
2. P. Gao, S.-M. Swei, Z. Yuan, A new piezodriven precision micropositioning stage utilizing flexure hinges. Nanotechnology **10**, 394–398 (1999)
3. A. Elmustafa, M.G. Lagally, Flexural-hinge guided motion nanopositioner stage for precision machining: finite element simulations. Precis. Eng. **25**, 77–81 (2001)
4. http://www.physikinstrumente.com
5. http://www.npoint.com
6. http://www.aerotech.com
7. S. Salapaka, A. Sebastian, J.P. Cleveland, M.V. Salapaka, High bandwidth nano-positioner: a robust control approach. Rev. Sci. Instrum. **73**, 3232–3241 (2002)
8. C.-W. Lee, S.-W. Kim, An ultraprecision stage for alignment of wafers in advanced microlithography. Precis. Eng. **21**, 113–122 (1997)
9. S. Akamine, T.R. Albrecht, M.J. Zdeblick, C.F. Quate, Microfabricated scanning tunneling microscope. IEEE Electron Device Lett. **10**, 490–492 (1989)
10. E. Meyer, H.J. Hug, R. Bennwitz, *Scanning Probe Microscopy: The Lab on a Tip* (Springer, New York, 2004)
11. J.J. Yao, S.C. Arney, N.C. MacDonald, Fabrication of high frequency two-dimensional nanoactuators for scanned probe devices. J. Microelectromech. Syst. **1**, 14–22 (1992)
12. Y. Xu, N.C. MacDonald, S.A. Miller, Integrated micro-scanning tunneling microscope. Appl. Phys. Lett. **67**, 2305–2307 (1995)

13. S.A. Miller, K.L. Turner, N.C. MacDonald, Microelectromechanical scanning probe instruments for array architectures. Rev. Sci. Instrum. **68**, 4155–4162 (1997)
14. S.C. Minne, J.D. Adams, G. Yaralioglu, S.R. Manalis, A. Atalar, C.F. Quate, Centimeter scale atomic force microscope imaging and lithography. Appl. Phys. Lett. **73**, 1742–1744 (1998)
15. S.C. Minne, G. Yaralioglu, S.R. Manalis, J.D. Adams, J. Zesch, A. Atalar, C.F. Quate, Automated parallel high-speed atomic force microscopy. Appl. Phys. Lett. **72**, 2340–2342 (1998)
16. D. Lange, T. Akiyama, C. Hagleitner, A. Tonin, H.R. Hidber, P. Niedermann, U. Staufer, N.F. de Rooij, O. Brand, H. Baltes, Parallel scanning AFM with on-chip circuitry in CMOS technology, in *Proceedings of IEEE MEMS* (IEEE, New York, 1999), pp. 447–452
17. T. Volden, M. Zimmermann, D. Lange, O. Brand, H. Baltes, Dynamics of CMOS-based thermally actuated cantilever arrays for force microscopy. Sensors Actuators A **115**, 516–522 (2004)
18. S. Hafizovic, D. Barrettino, T. Volden, J. Sedivy, K.U. Kirstein, O. Brand, A. Hierlemann, Single chip mechatronic microsystem for surface imaging and force response studies. Proc. Natl. Acad. Sci. **101**, 17011–17015 (2004)
19. A.G. Fowler, A.N. Laskovski, A.C. Hammond, S.O.R. Moheimani, A 2-DOF electrostatically actuated MEMS nanopositioner for on-chip AFM. J. Microelectromech. Syst. **21**, 771–773 (2012)
20. A. Mohammadi, A.G. Fowler, Y.K. Yong, S.O.R. Moheimani, A feedback controlled MEMS nanopositioner for on-chip high-speed AFM. J. Microelectromech. Syst. **23**, 610–619 (2014)
21. M. Maroufi, A. Bazaei, S.O.R. Moheimani, A high-bandwidth MEMS nanopositioner for on-chip AFM: design, characterization, and control. IEEE Trans. Control Syst. Technol. **23**, 504–512 (2015)
22. M. Maroufi, A.G. Fowler, A. Bazaei, S.O.R. Moheimani, High-stroke silicon-on-insulator MEMS nanopositioner: control design for non-raster scan atomic force microscopy. Rev. Sci. Instrum. **86**, 023705 (2015)
23. N. Sarkar, G. Lee, R.R. Mansour, CMOS-MEMS dynamic FM atomic force microscope, in *Proceedings of Transducers* (IEEE, New York, 2013), pp. 916–919
24. N. Sarkar, D. Strathearn, G. Lee, M. Olfat, R.R. Mansour, A 0.25 mm^3 atomic force microscope on-a-chip, in *Proceedings of IEEE MEMS* (IEEE, New York, 2015), pp. 732–735
25. E. Eleftheriou et al., Millipede – a MEMS-based scanning-probe data-storage system. IEEE Trans. Magn. **39**, 938–945 (2003)
26. Y. Ahn, T. Ono, M. Esahi, Si multiprobes integrated with lateral actuators for independent scanning probe applications. J. Micromech. Microeng. **15**, 1224–1229 (2005)
27. N. Sarkar, C. Baur, E. Stach, Z. Jandric, R. Stallcup, M. Ellis et al., Modular MEMS experimental platform for transmission electron microscopy, in *Proceedings of IEEE MEMS* (IEEE, New York, 2006), pp. 146–149
28. J.J. Gorman, Y.-S. Kim, A.E. Vladar, N.G. Dagalakis, Design of an on-chip microscale nanoassembly system. Int. J. Nanomanuf. **1**, 710–721 (2007)
29. X. Liu, J. Tong, Y. Sun, A millimeter-sized nanomanipulator with sub-nanometer positioning resolution and large force output. Smart Mater. Struct. **16**, 1742–1750 (2007)
30. S. Lu, D.A. Dikin, S. Zhang, F.T. Fisher, J. Lee, R.S. Ruoff, Realization of nanoscale resolution with a micromachined thermally actuated testing stage. Rev. Sci. Instrum. **75**, 2154–2162 (2004)
31. Y. Zhu, H.D. Espinosa, An electromechanical material testing system for in situ electron microscopy and applications. Proc. Natl. Acad. Sci. **102**, 14503–14508 (2005)
32. B. Pant, B.L. Allen, T. Zhu, K. Gall, O.N. Pierron, A versatile microelectromechanical system for nanomechanical testing. Appl. Phys. Lett. **98**, 053506 (2011)
33. S.T. Smith, D.G. Chetwynd, *Foundations of Ultra-Precision Mechanism Design* (CRC Press, Boca Raton, 1992)
34. S.T. Smith, *Flexures: Elements of Elastic Mechanisms* (Taylor & Francis, London, 2000)

35. R. Legtenberg, A.W. Groeneveld, M. Elwenspoek, Comb-drive actuators for large displacements. J. Micromech. Microeng. **6**, 320–329 (1996)
36. V.P. Jaecklin, C. Linder, N.F. de Rooij, J.M. Moret, R. Bischof, F. Rudolf, Novel polysilicon comb actuators for XY-stages, in *Proceedings of IEEE MEMS* (IEEE, New York, 1992), pp. 147–149
37. D. Kobayashi, T. Hirano, T. Furuhata, H. Fujita, An integrated lateral tunneling unit, in *Proceedings of IEEE MEMS* (IEEE, New York, 1992), pp. 214–219
38. M.I. Lutwyche, Y. Wada, Manufacture of micromechanical scanning tunnelling microscopes for observation of the tip apex in a transmission electron microscope. Sensors Actuators A **48**, 127–136 (1995)
39. C.S.B. Lee, S. Han, N.C. MacDonald, Single crystal silicon (SCS) XY-stage fabricated by DRIE and IR alignment, in *Proceedings of IEEE MEMS* (IEEE, New York, 2000), pp. 28–33
40. C.-H. Kim, Y.-K. Kim, Micro XY-stage using silicon on a glass substrate. J. Micromech. Microeng. **12**, 103–107 (2002)
41. J. Dong, D. Mukhopadhyay, P.M. Ferreira, Design, fabrication and testing of a silicon-on-insulator (SOI) MEMS parallel kinematics XY stage. J. Micromech. Microeng. **17**, 1154–1161 (2007)
42. Y.-S. Kim, J.-M. Yoo, S.H. Yang, Y.-M. Choi, N.G. Dagalakis, S.K. Gupta, Design, fabrication and testing of a serial kinematic MEMS XY stage for multifinger manipulation. J. Micromech. Microeng. **22**, 085029 (2012)
43. X. Liu, K. Kim, Y. Sun, A MEMS stage for 3-axis nanopositioning. J. Micromech. Microeng. **17**, 1796–1802 (2007)
44. J. Dong, P.M. Ferreira, Electrostatically actuated cantilever with SOI-MEMS parallel kinematic XY stage. J. Microelectromech. Syst. **18**, 641–651 (2009)
45. Y.S. Kim, N.G. Dagalakis, S.K. Gupta, Design of MEMS based three-axis motion stage by incorporating a nested structure. J. Micromech. Microeng. **24**, 075009 (2014)
46. D. Mukhopadhyay, J. Dong, E. Pengwang, P. Ferreira, A SOI-MEMS-based 3-DOF planar parallel-kinematics nanopositioning stage. Sens. Actuators A **147**, 340–351 (2008)
47. S.-C. Chen, M.L. Culpepper, Design of a six-axis micro-scale nanopositioner—μHexFlex. Precis. Eng. **30**, 314–324 (2006)
48. S.D. Senturia, *Microsystem Design* (Springer, New York, 2001)
49. C. Liu, *Foundations of MEMS* (Prentice Hall, Upper Saddle River, 2011)
50. D.J. Bell, T.J. Lu, N.A. Fleck, S.M. Spearing, MEMS actuators and sensors: observations on their performance and selection for purpose. J. Micromech. Microeng. **15**, S153–S164 (2005)
51. N.B. Hubbard, M.L. Culpepper, L.L. Howell, Actuators for micropositioners and nanopositioners. Appl. Mech. Rev. **59**, 324–334 (2006)
52. R.C. Batra, M. Porfiri, D. Spinello, Review of modeling electrostatically actuated microelectromechanical systems. Smart Mater. Struct. **16**, R23–R31 (2007)
53. Y.-M. Choi, J.J. Gorman, N.G. Dagalakis, S.H. Yang, Y.-S. Kim, J.-M. Yoo, A high-bandwidth electromagnetic MEMS motion stage for scanning applications. J. Micromech. Microeng. **22**, 105012 (2012)
54. J.D. Grade, H. Jerman, T.W. Kenny, Design of large deflection electrostatic actuators. J. Microelectromech. Syst. **12**, 335–343 (2003)
55. M. Olfatnia, S. Sood, J.J. Gorman, S. Awtar, Large stroke electrostatic comb-drive actuators enabled by a novel flexure mechanism. J. Microelectromech. Syst. **22**, 483–494 (2013)
56. T. Akiyama, U. Staufer, N.F. de Rooij, Atomic force microscopy using an integrated comb-shape electrostatic actuator for high-speed feedback motion. Appl. Phys. Lett. **76**, 3139–3141 (2000)
57. J.L.A. Yeh, C.-Y. Hui, N.C. Tien, Electrostatic model for an asymmetric combdrive. J. Microelectromech. Syst. **9**, 126–135 (2000)
58. M.S.C. Lu, G.K. Fedder, Position control of parallel-plate microactuators for probe-based data storage. J. Microelectromech. Syst. **13**, 759–769 (2004)
59. A.G. Onaran, M. Balantekin, W. Lee, W.L. Hughes, B.A. Buchine, R.O. Guldiken et al., A new atomic force microscope probe with force sensing integrated readout and active tip. Rev. Sci. Instrum. **77**, 023501 (2006)

60. E.C.M. Disseldorp, F.C. Tabak, A.J. Katan, M.B.S. Hesselberth, T.H. Oosterkamp, J.W.M. Frenken et al., MEMS-based high speed scanning probe microscopy. Rev. Sci. Instrum. **81**, 043702 (2010)
61. F.C. Tabak, E.C.M. Disseldorp, G.H. Wortel, A.J. Katan, M.B.S. Hesselberth, T.H. Oosterkamp et al., MEMS-based fast scanning probe microscopes. Ultramicroscopy **110**, 599–604 (2010)
62. E.S. Hung, S.D. Senturia, Extending the travel range of analog-tuned electrostatic actuators. J. Microelectromech. Syst. **8**(4), 497–505 (1999)
63. D.A. Horsley, N. Wongkomet, R. Horowitz, A.P. Pisano, Precision positioning using a microfabricated electrostatic actuator. IEEE Trans. Magn. **35**, 993–999 (1999)
64. Y. Sun, D. Piyabongkarn, A. Sezen, B.J. Nelson, R. Rajamani, A high-aspect-ratio two-axis electrostatic microactuator with extended travel range. Sens. Actuators A **102**, 49–60 (2002)
65. R. Cragun, L.L. Howell, Linear thermomechanical microactuators, in *Proceedings of ASME IMECE* (ASME, New York, 1999), pp. 181–188
66. L.L. Chu, Y.B. Gianchandani, A micromachined 2D positioner with electrothermal actuation and sub-nanometer capacitive sensing. J. Micromech. Microeng. **13**, 279–285 (2003)
67. N.B. Hubbard, L.L. Howell, Design and characterization of a dual-stage, thermally actuated nanopositioner. J. Micromech. Microeng. **15**, 1482–1493 (2005)
68. S. ergna, J.J. Gorman, N.G. Dagalakis, Design and modeling of thermally actuated MEMS nanopositioners, in *Proceedings of ASME International Mechanical Engineering Congress and Exposition* (ASME, New York, 2005), pp. 561–568
69. J.J. Gorman, Y.-S. Kim, N.G. Dagalakis, Control of MEMS nanopositioners with nano-scale resolution, in *Proceedings of ASME International Mechanical Engineering Congress and Exposition* (ASME, New York, 2006), pp. 151–159
70. R.K. Messenger, Q.T. Aten, T.W. McLain, L.L. Howell, Piezoresistive feedback control of a MEMS thermal actuator. J. Microelectromech. Syst. **18**, 1267–1278 (2009)
71. Y. Zhu, A. Bazaei, S.O.R. Moheimani, M.R. Yuce, Design, modeling, and control of a micromachined nanopositioner with integrated electrothermal actuation and sensing. J. Microelectromech. Syst. **20**, 711–719 (2011)
72. M. Rakotondrabe, A.G. Fowler, S.O.R. Moheimani, Control of a novel 2-DoF MEMS nanopositioner with electrothermal actuation and sensing. IEEE Trans. Control Syst. Technol. **22**, 1486–1497 (2014)
73. J.H. Comtois, V.M. Bright, M.W. Phipps, Thermal microactuators for surface-micromachining processes, in *Proceedings of SPIE 2642, Micromachined Devices and Components* (SPIE, Bellingham, 1995), pp. 10–21
74. D. Yan, A. Khajepour, R. Mansour, Design and modeling of a MEMS bidirectional vertical thermal actuator. J. Micromech. Microeng. **14**, 841–850 (2004)
75. T. Akiyama, U. Staufer, N.F. de Rooij, Fast driving technique for integrated thermal bimorph actuator toward high-throughput atomic-force microscopy. Rev. Sci. Instrum. **73**, 2643–2646 (2002)
76. G. Vitellaro, G. L'Episcopo, C. Trigona, B. Ando, S. Baglio, A compliant MEMS device for out-of-plane displacements with thermo-electric actuation. J. Microelectromech. Syst. **23**, 661–671 (2014)
77. P.J. Gilgunn, L. Jingwei, N. Sarkar, G.K. Fedder, CMOS-MEMS lateral electrothermal actuators. J. Microelectromech. Syst. **17**, 103–114 (2008)
78. D.O. Popa, B.H. Kang, J.T. Wen, H.E. Stephanou, G. Skidmore, A. Geisberger, Dynamic modeling and input shaping of thermal bimorph MEMS actuators, in *Proceedings of IEEE ICRA* (IEEE, New York, 2003), pp. 1470–1475
79. K.R. Oldham, J.S. Pulskamp, R.G. Polcawich, M. Dubey, Thin-Film PZT lateral actuators with extended stroke. J. Microelectromech. Syst. **17**, 890–899 (2008)
80. M.W. Pruessner, T.H. Stievater, W.S. Rabinovich, In-plane microelectromechanical resonator with integrated Fabry–Pérot cavity. Appl. Phys. Lett. **92**, 081101 (2008)
81. U. Krishnamoorthy, R.H. Olsson III, G.R. Bogart, M.S. Baker, D.W. Carr, T.P. Swiler, P.J. Clews, In-plane MEMS-based nano-g accelerometer with sub-wavelength optical resonant sensor. Sens. Actuator A **145–146**, 283–290 (2008)

82. L. Ji, Y. Zhu, S.O.R. Moheimani, M.R. Yuce, A micromachined 2DOF nanopositioner with integrated capacitive displacement sensor, in *Proceedings of IEEE Sensors* (IEEE, New York, 2010), pp. 1464–1467
83. Y. Zhu, S.O.R. Moheimani, M.R. Yuce, Simultaneous capacitive and electrothermal position sensing in a micromachined nanopositioner. IEEE Electron Device Lett. **32**, 1146–1148 (2011)
84. B. Koo, X. Zhang, J. Dong, S.M. Salapaka, P.M. Ferreira, A 2 degree-of-freedom SOI-MEMS translation stage with closed-loop positioning. J. Microelectromech. Syst. **21**, 13–22 (2012)
85. N. Yazdi, H. Kulah, K. Najafi, Precision readout circuits for capacitive microaccelerometers, in *Proceedings of IEEE Sensors* (IEEE, New York, 2004), pp. 28–31
86. J. Wu, G.K. Fedder, L.R. Carley, A low-noise low-offset capacitive sensing amplifier for a 50 μm/rt-Hz monolithic CMOS MEMS accelerometer. IEEE J. Solid-State Circuits **39**, 722–730 (2004)
87. J. Dong, P.M. Ferreira, Simultaneous actuation and displacement sensing for electrostatic drives. J. Micromech. Microeng. **18**, 035011 (2008)
88. S.I. Moore, S.O.R. Moheimani, Displacement measurement with a self-sensing MEMS electrostatic drive. J. Microelectromech. Syst. **23**, 511–513 (2014)
89. M.A. Lantz, G.K. Binnig, M. Despont, U. Drechsler, A micromechanical thermal displacement sensor with nanometre resolution. Nanotechnology **16**, 1089–1094 (2005)
90. Y. Zhu, A. Bazaei, S.O.R. Moheimani, M.R. Yuce, A micromachined nanopositioner with on-chip electrothermal actuation and sensing. IEEE Electron Device Lett. **31**, 1161–1163 (2010)
91. J. Chow, Y. Lai, Displacement sensing of a micro-electro-thermal actuator using a monolithically integrated thermal sensor. Sens. Actuators A **150**, 137–143 (2009)
92. A. Mohammadi, M.R. Yuce, S.O.R. Moheimani, A low-flicker-noise MEMS electrothermal displacement sensing technique. J. Microelectromech. Syst. **21**, 1279–1281 (2012)
93. A.A. Barlian, W.-T. Park, J.R. Mallon Jr., A.J. Rastegar, B.L. Pruitt, Review: semiconductor piezoresistance for microsystems. Proc. IEEE **97**, 513–552 (2009)
94. A. Bazaei, M. Maroufi, A. Mohammadi, S.O.R. Moheimani, Displacement sensing with silicon flexures in MEMS nanopositioners. J. Microelectromech. Syst. **23**, 502–504 (2014)
95. J. Ouyang, Y. Zhu, Z-shaped MEMS thermal actuators: piezoresistive self-sensing and preliminary results for feedback control. J. Microelectromech. Syst. **21**, 596–604 (2012)
96. Y. Ando, T. Ikehara, S. Matsumoto, Design, fabrication and testing of new comb actuators realizing three-dimensional continuous motions. Sens. Actuators A **97–98**, 579–586 (2002)
97. C.-H. Kim, H.-M. Jeong, J.-U. Jeon, Y.-K. Kim, Silicon micro XY-stage with a large area shuttle and no-etching holes for SPM-based data storage. J. Microelectromech. Syst. **12**, 470–478 (2003)
98. Y. Ando, Development of three-dimensional electrostatic stages for scanning probe microscope. Sensors Actuators A **114**, 285–291 (2004)
99. L. Gu, X. Li, H. Bao, B. Liu, Y. Wang, M. Liu et al., Single-wafer-processed nano-positioning XY-stages with trench-sidewall micromachining technology. J. Micromech. Microeng. **16**, 1349–1357 (2006)
100. X. Chen, D.-W. Lee, Integrated microactuation scanning probe microscopy system. J. Vac. Sci. Technol. B **27**, 1408–1412 (2009)
101. Y.-S. Kim, N.G. Dagalakis, S.K. Gupta, Creating large out-of-plane displacement electrothermal motion stage by incorporating beams with step features. J. Micromech. Microeng. **23**, 055008 (2013)
102. A.G. Fowler, A. Bazaei, S.O.R. Moheimani, Design and analysis of nonuniformly shaped heaters for improved MEMS-based electrothermal displacement sensing. J. Microelectromech. Syst. **22**, 687–694 (2013)
103. X. Zhang, B. Koo, S.M. Salapaka, J. Dong, P.M. Ferreira, Robust control of a MEMS probing device. IEEE/ASME Trans. Mechatron. **19**, 100–108 (2014)

Chapter 10
Control Issues of MEMS Nanopositioning Devices

Y. Zhu, S.O.R. Moheimani, M.R. Yuce, and A. Bazaei

Abstract In this chapter, the control issues of microelectromechanical system (MEMS) nanopositioning devices are introduced and discussed. The real-time feedback control of a novel micro-machined 1-degree-of-freedom (1-DoF) thermal nanopositioner with on-chip electrothermal position sensors is presented. The actuation works based on the thermal expansion of V-shaped silicon beams. The sensing mechanism works based on measuring the resistance difference between two electrically biased identical silicon beams. The resistance difference varies with displacement. The heat conductance of the sensor beams varies oppositely with the position of the movable stage, resulting in different beam temperatures and resistances. A pair of position sensors are operated in differential mode to reduce low-frequency drift. The micro-machined nanopositioner has a nonlinear static input–output characteristic. The electrothermal actuator has a dynamic range of 14.4 μm and the electrothermal sensor has a low drift of 8.9 nm over 2000 s. An open-loop controller is first designed and implemented. It is experimentally shown that uncertainties result in unacceptable positioning performance. Hence, feedback control is required for accurate positioning. The on-chip displacement sensor is able to provide high-resolution displacement control. Therefore, a real-time closed-loop feedback control system is designed using a proportional-integral (PI) controller together with the nonlinear compensator used for the open-loop control system. The closed-loop system provides acceptable and robust tracking resolution for a wide range of set point values. The step response results show a positioning resolution of 7.9 nm and a time constant of 1.6 ms in a 10 μm stroke. For triangular reference tracking, which is required in raster-scanned scanning probe microscopy (SPM), the steady-state tracking error has a standard deviation of 20 nm within a wide range of 10 μm.

Y. Zhu (✉) • S.O.R. Moheimani • M.R. Yuce • A. Bazaei
Griffith University, Nathan, QLD, Australia
e-mail: y.zhu@griffith.edu.au; yongyan.zhu@gmail.com

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_10

10.1 Introduction

High precision nanopositioners have been extensively used in many applications such as scanning tunneling microscopy (STM) [1], atomic force microscopy (AFM) [2], and emerging ultrahigh density probe storage systems [3–7]. These nanopositioners typically have high positioning accuracy with a large dynamic range and a wide bandwidth, enabling fast and robust closed-loop position control. Macro-scale nanopositioners can achieve nanometre-scale positioning resolution and accuracy, however they are relatively bulky and costly [8–11]. Microelectromechanical system (MEMS) nanopositioners have attracted significant interest recently because of their small size, low cost, fast dynamics, and the emergence of applications such as probe-based data storage [12, 13], and scanning probe microscopy [14–16]. Closed-loop feedback control of these positioners is highly desirable if a high degree of displacement precision is required, and such a control system needs an accurate source of position information [17–20]. However, many of the MEMS nanopositioners reported in the literature are not equipped with on-chip sensors due to the restrictions associated with micro-fabrication processes [21–23]. Thus, the in-plane movements are often measured by laser reflectance microscopes [24, 25] or optical microscopes [26], making the footprint of the entire system fairly large. There are several exceptions in the literature, e.g., embedded on-chip capacitive and piezo-resistive displacement sensors were integrated in thermally and electrostatically actuated positioners in [27–30]. Nevertheless, only open-loop results were obtained and complex fabrication processes were required for electrical insulation between electrical heating and sensing circuits in [27] and for sidewall implementation of the piezo-resistors in [29]. In addition, limited sensing resolutions were achieved in [28] and [29]. Recently, a thermal sensing scheme was used in a probe-based storage device [31, 32]. Micro-heaters were used to measure the motion of a MEMS micro-scanner with resolution of less than 1 nm. Compared to comb capacitive sensors, thermal sensors are much more compact and can be easily integrated with actuators in a MEMS device [33–35]. In [36, 37], off-chip electromagnetic coil actuators were adopted for scanner actuation along with the thermal sensing approach reported in [31, 32], where a complex mass-balanced structure was designed for vibration resistance purposes.

In this chapter, a novel electrothermal position sensor is integrated with an electrothermal actuator in the same MEMS chip without the need for inclusion of extra electrical insulation fabrication process as reported in [27], or assembling two chips as reported in [36, 37]. Although the electrothermal actuators consume more power compared with other MEMS actuators, for example, electrostatic actuator and piezoelectric actuators, they do offer certain advantages. They can operate under low voltages, generate large forces, and enjoy a high degree of vibration resistance due to their stiff structures [38].

This chapter introduces a MEMS nanopositioner with on-chip electrothermal actuation and electrothermal sensing, which was micro-fabricated in a low-cost bulk silicon process. Measurement results show that the positioner has a dynamic

range of 14.4 μm and an open-loop bandwidth of 101 Hz. To reduce the low-frequency thermal drift, the sensors are operated in a pair and measured by a differential circuitry. The sensor was measured to have a small low-frequency drift of 8.9 nm over 2000 s. The on-chip displacement sensing enables a feedback control capability. A model of the positioner is derived and a proportional-integral (PI) feedback controller is implemented digitally in a dSPACE rapid prototyping system to investigate the closed-loop performance of the positioner. Open-loop and closed-loop performance of the system in tracking a step displacement were investigated. The closed-loop step response results show a steady-state RMS positioning error of 8.6 nm, while the corresponding value of the open-loop seek operation can reach up to 1180 nm. High tracking performance was also achieved for a 10 Hz triangular reference by using a 2-DoF feedback control system consisting of a PI controller and a prefilter.

The chapter is organized as follows: In Sect. 10.2, the electrothermal positioning scheme and its modeling are introduced. Micro-fabrication process and design parameters are described in Sect. 10.3. Section 10.4 presents the experimental setup and system characterization of the positioner. Open-loop static and dynamic responses are presented to verify that the positioner behaves as predicted. Section 10.5 presents the details of the open-loop and closed-loop control design. The tracking performances of step and triangle references are presented in this section. Section 10.6 concludes this chapter.

10.2 Design and Modeling of Micro-Machined Nanopositioner

10.2.1 Design

The proposed nanopositioner consists of electrothermal actuator and sensor on a single silicon chip. On-chip actuation and sensing enable a compact positioning system with small footprint and high resolution. Figure 10.1 illustrates the conceptual schematic view of the nanopositioner. The positioner stage is actuated by a chevron electrothermal actuator. To move stage, dc current is applied across the heaters to heat up the silicon beams. Compared to other actuation methods, for example, electrostatic actuation and piezoelectric actuation, the electrothermal actuators are simple to implement with low actuation voltage and large force. Thanks to the high stiffness of the structures, complex mass-balanced structures as proposed in [37] are not required to resist vibration. As shown in Fig. 10.1, a pair of thin hot arms are designed to have an angle of 0.57° with respect to each other in the V-shaped electrothermal actuator. The operating voltage range is typically from 0 to 15 V in chevron actuators, which depends on the beam geometries. The displacement is proportional to the square of actuation voltage and the largest stroke of the actuator is limited by the buckling of the hot arms at high temperatures (>700 °C) [25].

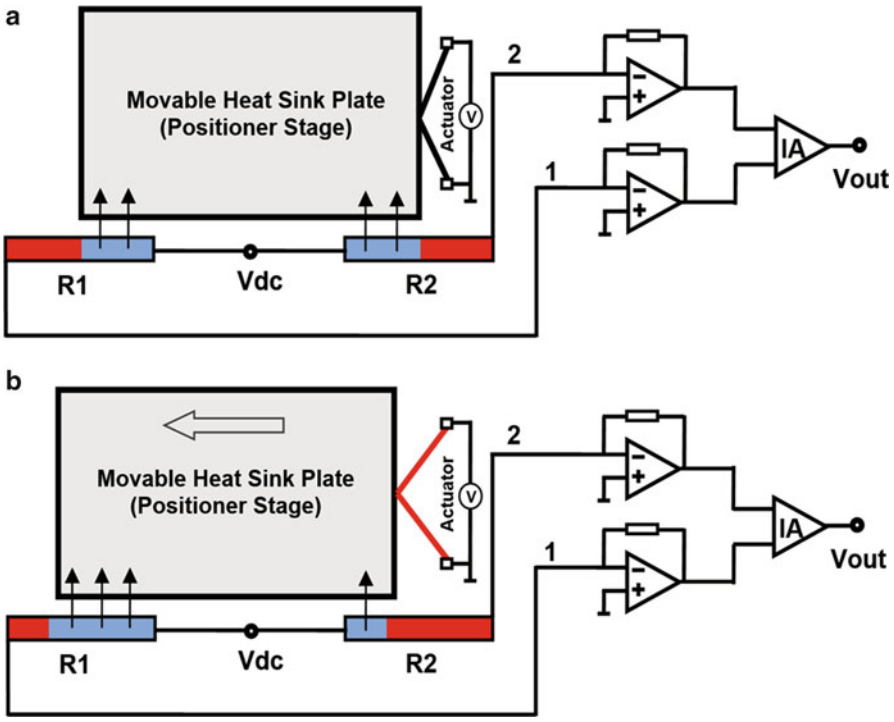


Fig. 10.1 Schematic diagram of the nanopositioner with electrothermal actuator, electrothermal position sensor, and differential amplifier: (**a**) before actuation; (**b**) after actuation [39]

The position sensors are based on electrothermal principle. Two beam-shaped resistive heaters R_1 and R_2 are made from n-doped silicon (2 μm wide, 100 μm long, and 25 μm thick). The positioner stage (rectangular plate) is placed close to the beam heaters with a 2 μm air gap. Initially, the positioner stage is at the rest position before applying voltage across the actuator, as illustrated in Fig. 10.1a. At this position, the two edges of the sink plate are precisely aligned to the center of the two thermal resistive sensors R_1 and R_2. To heat up the sensors, a dc voltage source V_{dc} is applied to both sensors (heaters). Consequently, the heat generated in the resistive heaters is conducted through the 2 μm air gap to the heat sink plate (positioner stage). At the rest position in Fig. 10.1a, the ambient condition for both sensors is identical; therefore, the dissipated heat fluxed out of the sensors is same. Thus, it equalizes the temperature and resistance of the two sensors.

In Fig. 10.1b, a voltage is applied on the actuator beams to move the positioner stage towards left. As a consequence, the ambient condition is changed to be different. The heat fluxes dissipated out of the both sensors become different, resulting in a corresponding difference in the temperatures of both sensors. Also, the resistance is the function of temperature in silicon material. Therefore the resistance of sensor R_1 and sensor R_2 will be different. Thus, the positioner stage's displacement can be

sensed by measuring the resistance changes between the two sensors. With a fixed dc voltage V_{dc} applied to both sensors, the differential changes of the resistance result in current variations in the beam resistors because of Ohm's law. The currents are then converted to an output voltage using a pair of trans-impedance amplifiers and an instrumentation amplifier. To suppress the common-mode noise, the gains of these two trans-impedance amplifiers must be well matched by adjusting the feedback resistance of the trans-impedance amplifiers. Employing the differential topology allows the sensor output to be immune from undesirable drift effects due to changes in ambient temperature or aging effects.

10.2.2 Modeling

This section presents a large-signal analysis for the anticipation of temperature and resistance values of the heaters in terms of a pre-specified displacement and bias voltage. We adopt a lumped parameter approach and static (time-invariant) conditions to simplify the analysis [40]. Each sensing resistor is regarded as an infinitesimal solid element with no internal thermal convection. With this assumption, the resistor temperature is presented by a scalar variable rather than a spatial distribution, which would require solving a complicated boundary value PDE problem. In addition, using this assumption, we can combine a number of heat transfer coefficients between the object and various parts in its surrounding such as heat sink, air gap, and conducting pads, into an overall thermal resistance coefficient. Applying Newton's law of cooling, the transfer of the heat from the object to air is mainly governed by thermal conduction in a thin air layer beside the object instead of thermal convection. Although thermal convection can participate in the transfer of heat outside the foregoing thin layer, its overall affect in stationary conditions can be modeled as an equivalent thermal resistance between the object and air. In this way, each sensing resistor has a uniform temperature distribution. We now proceed to a lumped parameter analysis for temperature prediction of heaters.

Assuming a constant temperature coefficient $\alpha(1/°C)$, the electrical resistance R of an infinitesimal resistor at temperature T is described by $R_o[1 + \alpha(T - T_o)]$, where T_o and R_o (Ω) describe the ambient temperature and the resistance value at room temperature, respectively. Let the overall thermal resistance between the resistor and the surrounding area, including the thermal conduction through its electrical wiring to a voltage source, be $K^{-1}(°C/W)$. Assuming a constant bias voltage V across the resistor and a stationary condition, the electrical power delivered to the resistor should be equal to the heat power delivered by the resistor to the ambient area:

$$\frac{V^2}{R_o[1 + \alpha(T - T_o)]} = K(T - T_o) \tag{10.1}$$

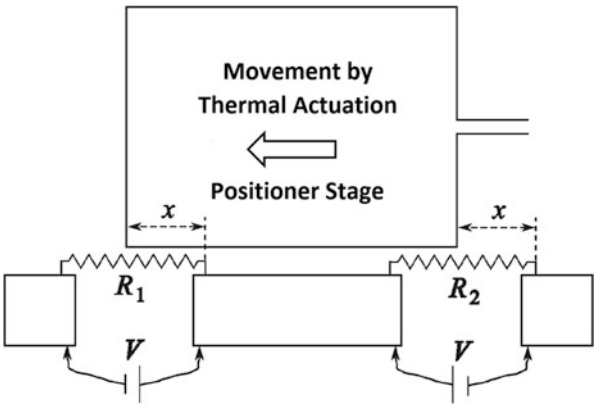


Fig. 10.2 Schematic diagram of the biased electrothermal position sensors [39]

Thanks to the simplifying assumption that constant α does not vary with the temperature, we can obtain a closed-form solution for the temperature T from Eq. (10.1). Hence, the temperature and electrical resistance of the resistor at bias voltage V are determined by the following expressions:

$$R = R_{\mathrm{o}} \frac{1 + \sqrt{1 + \frac{4\alpha V^2}{R_{\mathrm{o}} K}}}{2} \tag{10.2}$$

$$T = T_{\mathrm{o}} + \frac{\sqrt{1 + \frac{4\alpha V^2}{R_{\mathrm{o}} K}} - 1}{2\alpha} \tag{10.3}$$

Before applying Eqs. (10.2) and (10.3) to the heaters in Fig. 10.2, we need to introduce individual heat resistances ${K_1}^{-1}$ and ${K_2}^{-1}$ to the ambient area for each electric resistors R_1 and R_2, respectively. A horizontal displacement of the heat sink plate in the left direction will increase the heat conductance of the left heater, whilst reducing that of the right heater. These monotonic variations of the heat conductance values with displacement can roughly be modeled by the following linear relationships:

$$\begin{aligned} K_1(\tilde{x}) &= K_{\min} + (K_{\max} - K_{\min})\,\tilde{x}, \\ K_2(\tilde{x}) &= K_{\min} + (K_{\max} - K_{\min})\,(1 - \tilde{x}) \end{aligned} \tag{10.4}$$

where $\tilde{x} \equiv x/L$ describes a dimensionless position variable, L is the heater length, and ${K_{\min}}^{-1}$ and ${K_{\max}}^{-1}$ represent the maximum and minimum overall heat resistance of each heater, respectively. We can now use Eqs. (10.2) and (10.3) to predict the stationary values of temperature and resistance for each heater:

Table 10.1 Electrothermal sensor parameters and values for simulation

Parameters	Values
$\alpha(\mathrm{K}^{-1})$	3.7822×10^{-4}
$R_{\mathrm{o}}(\Omega)$	390
$T_{\mathrm{o}}(^{\circ}\mathrm{C})$	27
$K_{\min}(\mathrm{W/K})$	8×10^{-5}
$K_{\max}(\mathrm{W/K})$	16×10^{-5}
$V(\mathrm{V})$	6

$$
\begin{aligned}
R_i\left(\tilde{x}, V\right) &= R_{\mathrm{o}} \frac{1+\sqrt{1+\frac{4\alpha V^2}{R_{\mathrm{o}} K_i(\tilde{x})}}}{2}, \\
T_i\left(\tilde{x}, V\right) &= T_{\mathrm{o}} + \frac{\sqrt{1+\frac{4\alpha V^2}{R_{\mathrm{o}} K_i(\tilde{x})}}-1}{2\alpha}, i \in \{1,2\}
\end{aligned}
\tag{10.5}
$$

where R_{o} is value of unbiased heater electrical resistance at room temperature. When the heat sink displacement is zero($\tilde{x} = 0.5$), we have assumed an identical heat conductance $K_1 = K_2 = (K_{\max} + K_{\min})/2$ for the heaters. Hence, the sensing resistors are also identical when the displacement is zero. In this way, the output voltage in Fig. 10.1 is proportional to $R_1^{-1} - R_2^{-1}$. This result is used in the following simulation to convert the amount of heat sink displacement into a dimensionless variable $R_{\mathrm{o}}\left(R_1^{-1} - R_2^{-1}\right)$, which is proportional to the output voltage in Fig. 10.1.

Considering the values reported in Table 10.1 for model parameters and applying Eqs. (10.4) and (10.5), the stationary temperature, resistance, and sensor output values as a function of heat sink displacement are obtained for each sensing heater as depicted in Fig. 10.3a–c, respectively. In spite of the nonlinear variation of the individual resistance values with displacement, the sensor output demonstrates a highly linear relationship with the displacement.

10.3 Microfabrication

The device was fabricated in a commercial silicon-on-insulator (SOI) Micromachining process in a MEMS foundry (MEMSCAP) [41]. This process allows fabrication of silicon structures with a fixed 25 μm of thickness and a minimum 2 μm of width and gap. The final device structure is illustrated in Fig. 10.4, and the fabrication process is briefly described as follows:

1. Microfabrication starts with a highly doped n-type 25 μm silicon device layer, on which metal pads (20 nm Cr and 500 nm Au) are deposited by e-beam evaporation and patterned through lift-off process. The metal pads are to provide ohmic contact for probes and/or bonding wires.
2. Silicon is lithographically patterned to define the structures. Then deep reactive ion etch (DRIE) is applied from the front side of the wafer to define both the anchored and movable features of the structure.

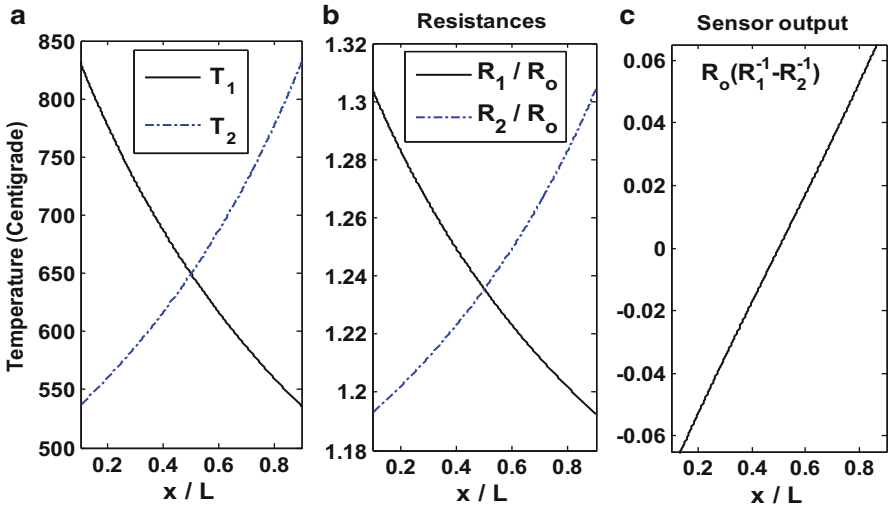


Fig. 10.3 Simulated steady-state values versus displacement for (**a**) resistor temperatures, (**b**) resistance values, and (**c**) normalized sensor output [39]

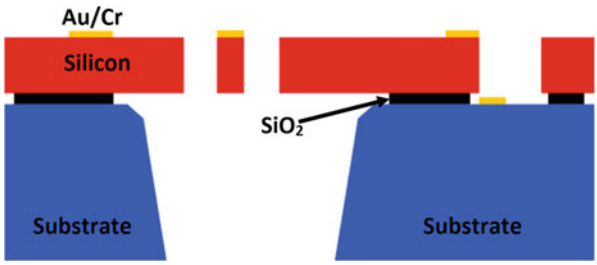


Fig. 10.4 SOIMUMPs microfabrication structures from MEMSCAP [41]

3. A protective polyimide layer is applied to the top surface of the silicon layer to protect the silicon structures during the following etching process.
4. Another DRIE is applied from the back side of the wafer to etch completely through the substrate in defined areas, stopping on the buried oxide layer. Then a trench underneath the movable structures is created.
5. The exposed buried oxide is removed using a wet oxide etch process (HF etching).
6. The polyimide coat on the front side is removed by oxygen plasma, thereby allowing the movable structure to be fully released. The front side oxide layer is removed using a vapor HF process. Then a large contact metal pad (i.e. electrical grounded contact) is patterned on the substrate to reduce parasitics.

The scanning electron microscope (SEM) image of the whole device and a close-up section of it are provided in Fig. 10.5.

To reduce the thermal coupling effects from the electro-thermal actuators to thermal sensors, a number of holes were embedded in the center shuttle between

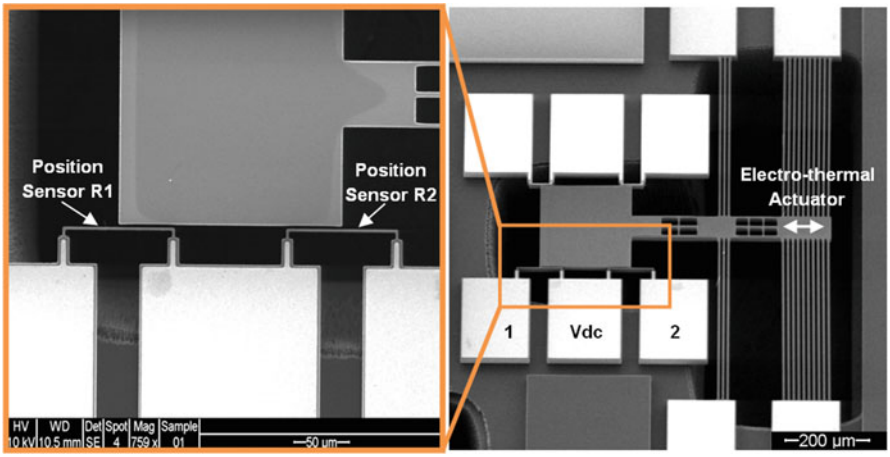


Fig. 10.5 SEM images of the micro-machined nanopositioner. The close-up section shows the details of a pair of electrothermal sensors, and positioner stage as a heat sink for the sensors [39]

Table 10.2 Design parameters and values

Parameters	Values
Thermal actuator beam width (μm)	5
Thermal actuator beam length (μm)	2×400
Thermal actuator beam angle (degree)	0.57
Thermal actuator beam quantity (pair)	10
Suspension beam width (μm)	5
Suspension beam length (μm)	400
Suspension beam angle (degree)	0
Suspension beam quantity (pair)	3
Thermal sensor beam width (μm)	2
Thermal sensor beam length (μm)	100
Gap between thermal sensor and heat sink plate (μm)	2
Heat sink plate dimension (μm × μm)	200×200
Device thickness (μm)	25
Measured resistance of thermal sensor R_1 at room temperature (Ω)	388.9
Measured resistance of thermal sensor R_2 at room temperature (Ω)	391.6
Measured resistance of electro-thermal actuator at room temperature (Ω)	116

actuators and the heat sink plate. The holes are expected to improve the thermal convection, thereby thermally insulating the heat sink plate from actuators. The suspension beams help the thermal insulation as well, due to the heat transfers to the substrate through the suspension beams and its anchors. Table 10.2 summarizes the device design parameters and measured beam resistance.

10.4 Experimental Characterization

The displacement of the nanopositioner was calibrated using a Polytec^{TM} Planar Motion Analyzer (PMA-400, Polytec GmbH). To accurately characterize the in-plane displacement of the positioner stage, digital image capturing and analysis methods were utilized. The PMA has an ordinary speed camera with a limited exposure time of 82 ms, however it can measure periodic processes at frequencies as high as 1 MHz by using the stroboscopic principle [42]. Step response is essential to characterize the dynamic performance of the positioner. Time constant and displacement can be obtained through step response measurement. Therefore, step voltages were applied to the actuator periodically, and the PMA recorded the displacement time response in each period, as shown in Fig. 10.6. The step time response of the actuator has been characterized using different actuation voltages of 8.5 and 9.0 V to exam the step displacement and time constant. As illustrated in Fig. 10.7, after 20 ms of actuation, the positioner stage reached the stable positions of 13.6 and 14.4 μm with 8.5 and 9.0 V step actuation voltages, respectively. The time constant is approximately 1.6 ms for both cases.

To control the positioner stage with high resolution, the static characterization of the actuator and the sensor must be implemented accurately. Particularly, two calibration curves are required: displacement versus actuation voltage curve for the actuator and output voltage versus displacement curve for the sensor. Accurate displacement measurement is crucial for the above two calibration curves. Therefore, PMA is utilized to measure the in-plane displacement. The static displacement versus actuation voltage measured data with a fitted fifth order polynomial is illustrated in Fig. 10.8a. With the actuation voltage of 9 V, the electrothermal

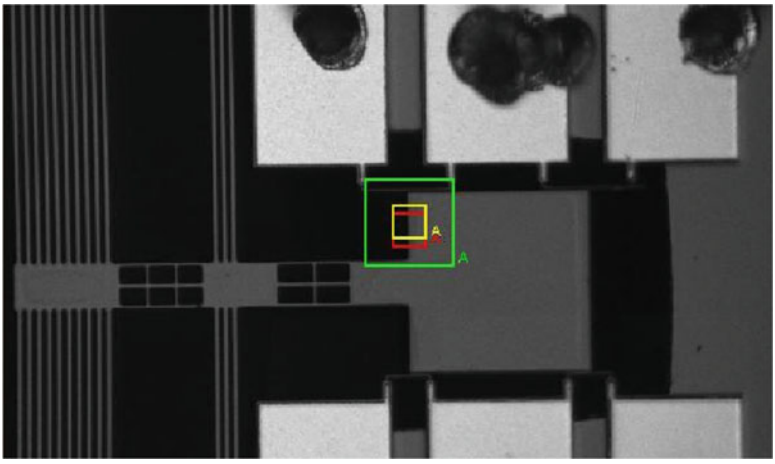


Fig. 10.6 Displacement measurement using Polytec^{TM} Planar Motion Analyzer (PMA-400, Polytec GmbH)

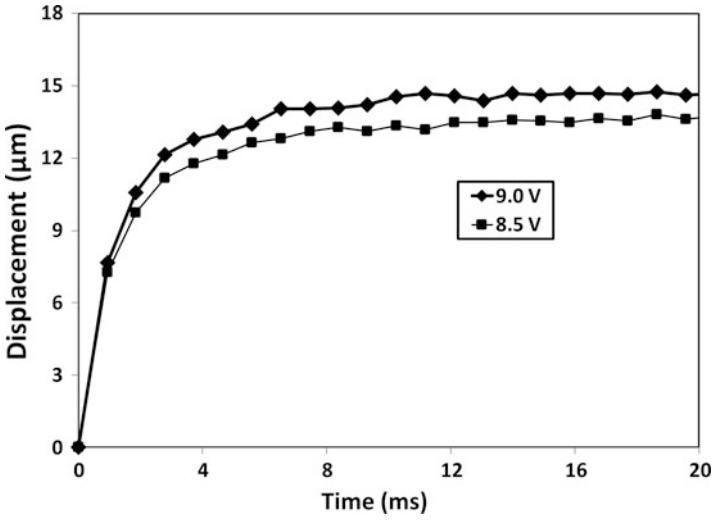


Fig. 10.7 Time response of the electrothermal actuator with step actuation voltages of 8.5 and 9.0 V

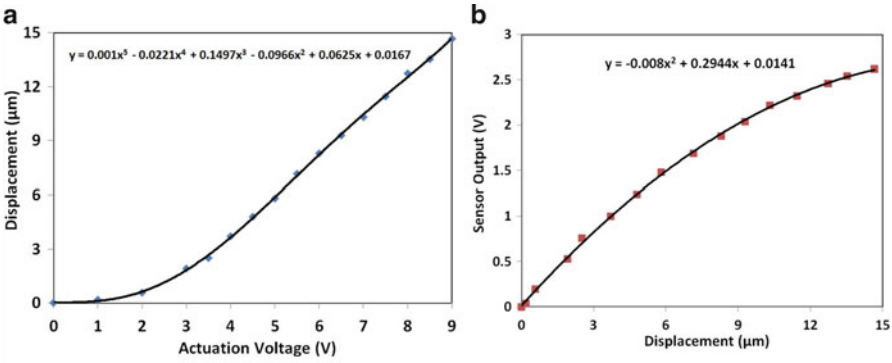


Fig. 10.8 Experimental calibration curves: (**a**) displacement versus actuation voltage for actuator; (**b**) output voltage versus displacement for sensor

actuator can achieve a maximum displacement of 14.4 μm, which is matched well with the step response measurement in Fig. 10.7. Due to the nature of the thermal expansion, the actuator can only move forward in one direction after application of a voltage. However, the actuator can be biased to an initial position by a dc offset voltage, such that it can move forward or backward by increasing or decreasing the applied voltage. For instance, for a 10 μm travel range with bidirectional movement, the actuator can be actuated by a 4.5 V dc voltage to the middle position (5 μm) of the desired range.

To calibrate the position sensor, the instrumentation amplifier outputs were recorded at every actuation voltage, as illustrated in Fig 10.8b. The sensors were

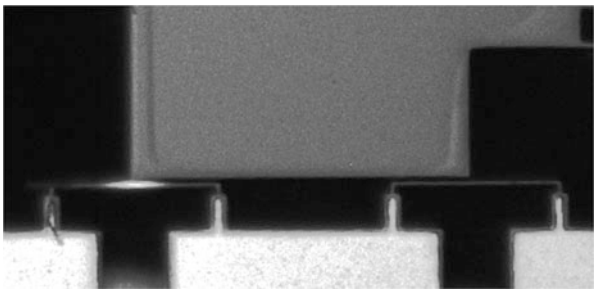


Fig. 10.9 Optical micrograph of the position sensor's illumination with an applied voltage of 10 V to the left silicon beam. No voltage was applied to the right sensor for comparison purpose [39]

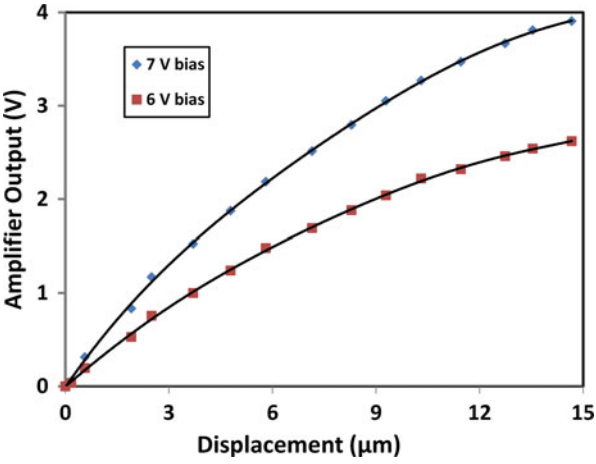


Fig. 10.10 Output voltage versus displacement with different sensor bias voltages of 6 and 7 V [39]

biased with 6 V and the instrumentation amplifier gain was set at 90.3 V/V. With the bias voltage of 6 V, a power of 60 mW is consumed in each sensor and the sensitivity is 0.27 mV/nm with a displacement up to 6 μm. As shown in Fig. 10.8b, the sensitivity drops when the displacement is more than 6 μm. A possible reason is that the temperature distribution is not uniform in the thermal sensor beams. Because of the heat loss at the anchors of the sensor beams, the center of the beam is hotter than its two ends, as described in [43].

To visualize the temperature distribution, a 10 V dc voltage was applied to the left sensor in Fig. 10.9. Visible light was observed in the center of the silicon beam while the two ends of the sensor beams have no light. The right sensor is dark because no voltage was applied to it for the comparison purpose. It was observed that 10 V is safe for the silicon beam without damage. However, the bias voltage should be far less than 10 V for the sensor's reliability reason. Therefore, the sensor is calibrated with 6 V bias voltage.

It will be of interest to investigate the effect on the sensor sensitivity from the bias voltage. The relationships between positioner displacement and sensor output at various sensor biases are illustrated in Fig. 10.10. The sensitivity can be increased by increasing the sensor bias voltage. However, the nonlinear nature of the response could not be eliminated by increasing the bias voltage.

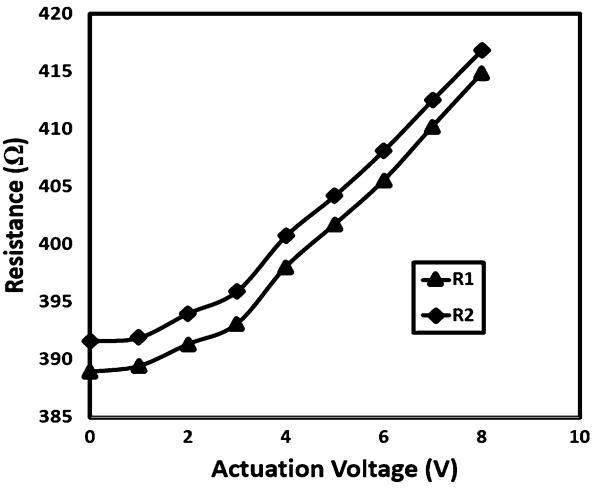


Fig. 10.11 Measured resistance versus voltage with no bias voltage applied to the sensor pair

As actuator and sensor are on a single silicon chip, we cannot assume that the displacement does not have any effect on the temperature of the area around the sensor resistors. To investigate the thermal coupling from actuator to sensor, the sensor resistances were measured using multimeter without applying an external bias voltage. As shown in Fig. 10.11, it is observed that the un-biased sensor resistances increase by about 6.4 % for a 12 μm displacement, just due to transfer of heat from actuator to sensor. The difference between the two curves in Fig. 10.11 is because of the variations and errors in beam dimensions during microfabrication.

Dynamic characterization is crucial for control implementation. HP35670A spectrum analyzer was utilized to characterize the positioner's dynamic performance. To characterize the frequency response, the nanopositioner can be regarded as an input–output system, which is connected to the spectrum analyzer. The input actuation voltage versus sensor output voltage curve can be obtained from the measured data in Fig. 10.8. The fitted fifth order polynomial calibration curve is illustrated in Fig. 10.12. To drive the actuator at high frequency, a dc voltage of 4.5 V superimposed with an ac voltage of 1 V was applied to the electrothermal actuator. Swept sinusoidal measurements were carried out from 1 Hz to 51.2 kHz by applying the input voltage to the actuator and sensing from the on-chip thermal sensors. The frequency response, illustrated in Fig. 10.13, shows that the open-loop bandwidth of the positioner is 101 Hz, which is relatively low compared to the thermal structures fabricated by PolyMUMPs [44]. The reason is because of the fact that the heat transfer to surrounding air in SOIMUMPs is much slower than that to the substrate in PolyMUMPs. However, the bandwidth of 101 Hz is satisfactory for some positioning applications.

Figure 10.14 shows the open-loop sensor drift when the positioner is at rest, which was measured at the output of the instrumentation amplifier over a period of 2000 s under normal laboratory conditions. Low-frequency drift can be calculated

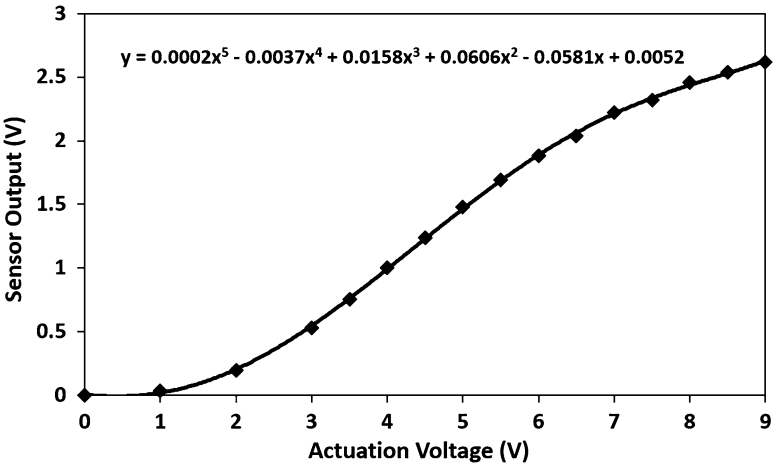


Fig. 10.12 Experimental calibration curve: actuation voltage versus sensor output voltage for the overall system including actuator and sensor

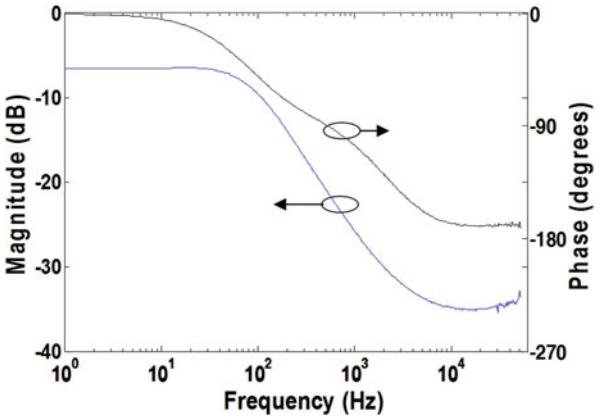


Fig. 10.13 Measured open-loop frequency response of the MEMS nanopositioner using HP35670A spectrum analyzer

by the moving average of the measured data. Thanks to the differential sensing of the sensor pair, the open-loop amplifier output has a low drift of 2.4 mV over 2000 s, which corresponds to 8.9 nm displacement.

10.5 Control Implementation

Having known the static and dynamic parameters of electrothermal actuators and sensors, proper control strategies can be implemented to achieve high positioning resolution and stability for the MEMS nanopositioner. In this section, open-loop and closed-loop position control strategies are investigated to control the MEMS positioner to have a positioning resolution in nanometre range. A dSPACE-1103 rapid

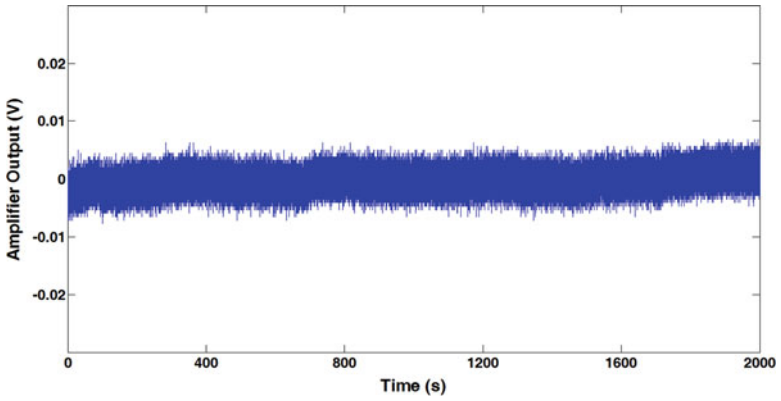


Fig. 10.14 Experimental results of sensor drift (open-loop) over 2000 s

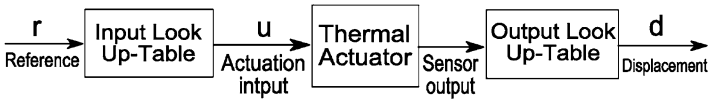


Fig. 10.15 Block diagram of the open-loop control system including input and output look up tables, reference, electrothermal actuator and sensor [39]

prototyping system was used for real-time implementation of the controllers and data acquisition. The dSPACE system is fully programmable from the Simulink® block diagram in Matlab environment.

10.5.1 Open-Loop Control

The block diagram of the open-loop control system is shown in Fig. 10.15. Accurate look-up-tables are required for fast computation of static nonlinear input–output mappings. For a specific input value to the block, it generates an output value based on interpolation or extrapolation among the recorded data values. The look-up-table at the input of the electro-thermal actuator converts the reference signal in micrometers to an appropriate actuation voltage, as illustrated in Fig. 10.16a, which was obtained using PMA as described in Sect. 10.5. However, the number and precision of measured data points are not enough to provide a smooth curve. Therefore, a fifth order polynomial was fitted to the measured data in Fig. 10.16a to generate the data for the input look-up-table. In a similar way, the look-up-table at the output of the electrothermal actuator was obtained to convert the sensor output voltage to displacement in micrometers, as shown in Fig. 10.16b. For a sensor bias voltage of 7 V, a third order polynomial was fitted to the measured data points in Fig. 10.10 to generate smooth and monotonic data for the output look-up-table.

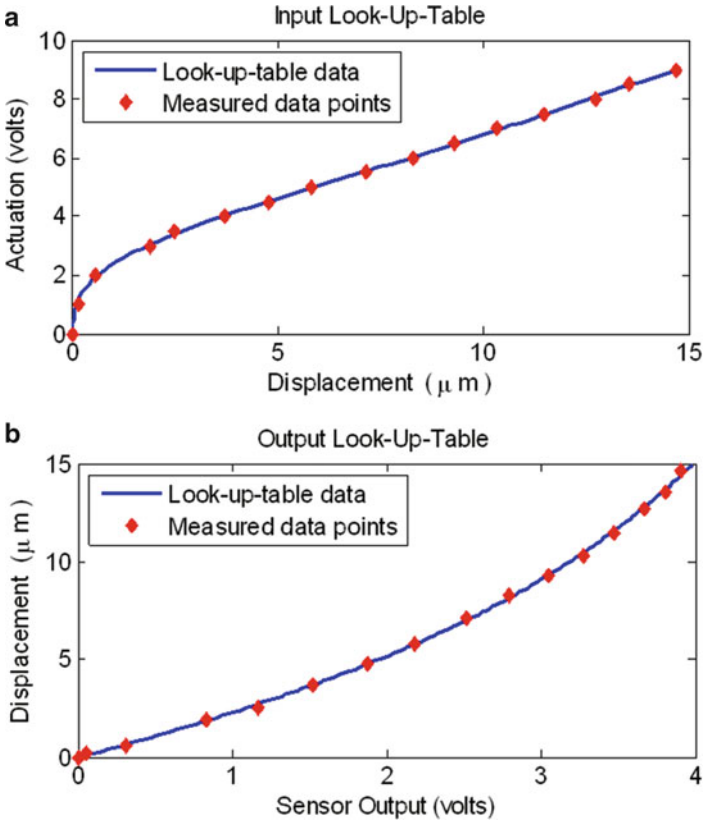


Fig. 10.16 The data used in input and output look-up-tables [39]

To avoid the infinite slope of the curve around zero displacement in Fig. 10.16a, the origin for displacement was shifted to +1 micrometer. Then both input and output look-up-tables were input to Simulink to control the electrothermal actuator.

To evaluate the control capability of the open-loop control method, a staircase reference signal is utilized as shown in Fig. 10.17. The reference voltage increased from 0 to 10 V with 2.5 V steps and 50 ms time intervals, and then decreased to 0 V in same manner. Unfortunately, it can be seen in Fig. 10.17 that the open-loop controller cannot provide an acceptable tracking performance for staircase displacement. A maximum displacement error of 0.96 μm (9.6 % error) was observed at 10 μm stroke. This is due to the plant uncertainties and the drift in the sensor and electronic circuits. Plant herein refers to the combination of the electrothermal MEMS device and electronic circuits for actuation and sensing. Therefore, feedback control strategies have to be considered to compensate the uncertainties and errors for accurate positioning.

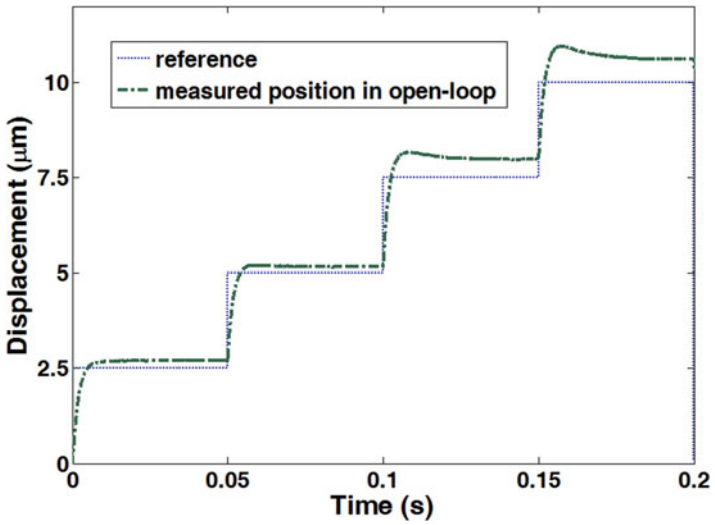


Fig. 10.17 Tracking performance of open-loop control method for a staircase reference. Large over-shoot errors are observed

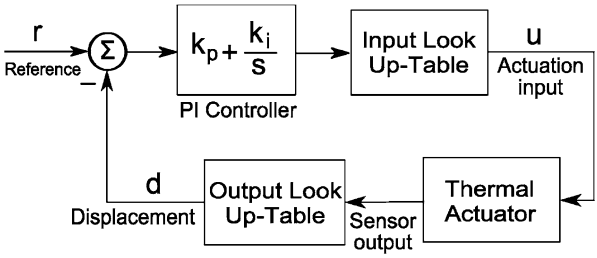


Fig. 10.18 Block diagram of one-degree-of-freedom (1DoF) PI control system including input and output look up tables, reference, electrothermal actuator and sensor, PI controller [39]

10.5.2 Closed-Loop Control

For closed-loop control, we incorporated a proportional-integral controller (PI controller) in addition to the nonlinear mappings, as illustrated in Fig. 10.18. The integration part in the PI controller provides a closed-loop unity low frequency gain form reference to displacement for set point tracking and robustness to uncertainties and disturbances. The proportional part in the PI controller is utilized to reduce the overshoot in step response. The integral gain and proportional gain were determined by experimental measurements. With an integral gain of $k_i = 700$ and a proportional gain of $k_p = 1.3$, the positioning performance is significantly improved compared to open-loop case, as shown in Fig. 10.19. Based on this control scheme, a controllable desired response of 2.5 μm steps over a 10 μm range was obtained with a maximum time constant of 1.4 ms. As illustrated in Fig. 10.19, the overshoot issue in open-loop control was also eliminated.

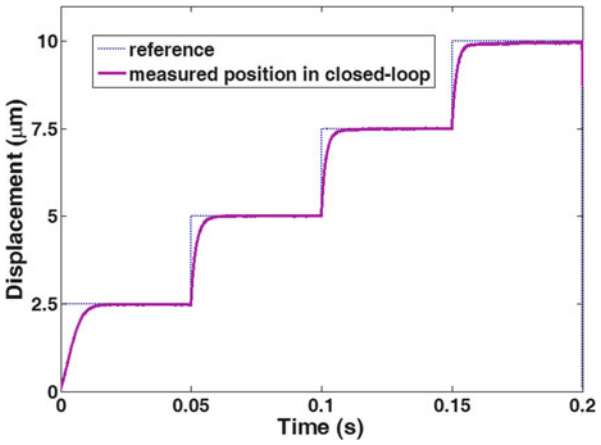


Fig. 10.19 Tracking performance of 1DoF closed-loop control method for a staircase reference

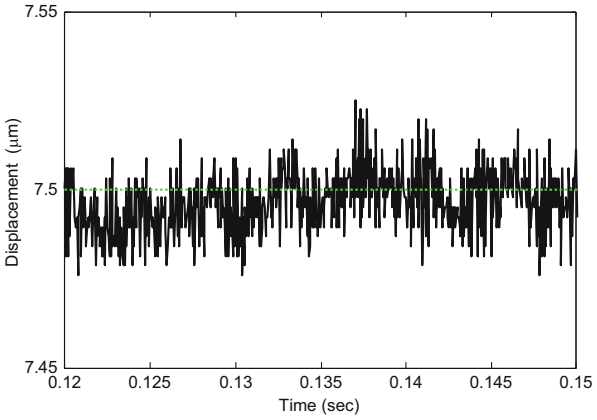


Fig. 10.20 Close-up view of the 7.5 μm closed-loop seek operation in Fig. 10.19. RMS positioning error of 8.6 nm is observed [39]

Figure 10.20 is the close-up view of the 7.5 μm closed-loop seek operation in Fig. 10.19, which indicates a steady-state RMS positioning error of 8.6 nm. As a comparison, the steady-state RMS positioning error of a similar open-loop seek operation can reach up to 1180 nm.

To further investigate the tracking capability of the nanopositioner, a 10 Hz triangular reference covering a 10 μm range was applied. The one-degree-of-freedom (1DoF) PI control system in Fig. 10.18 yielded an acceptable control performance, as shown in Fig. 10.21 (labeled "No prefilter"). With an integral gain of $k_i = 2000$ and a proportional gain of $k_p = 4$, the steady-state tracking error has a standard deviation of 90 nm, which is less than 1 % of the 10 μm full range.

The control performance can be further improved with a 2DoF control system consisting of a PI controller and a prefilter. The look-up-tables provide the nonlinear mappings described earlier. To provide better stability margins we used lower

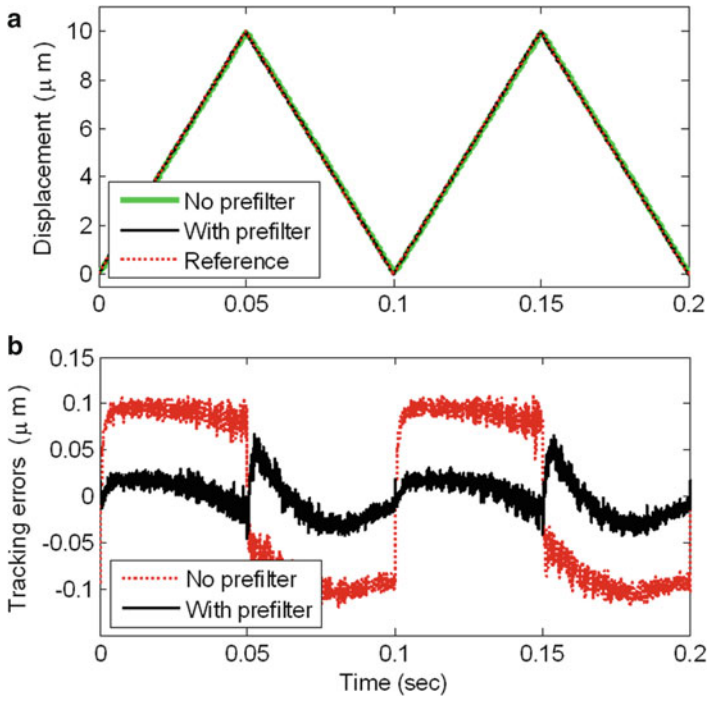


Fig. 10.21 Performance of 1DoF (no prefilter) and 2DoF (with prefilter) control systems in tracking of a triangular reference. RMS positioning error of 90 nm is observed [39]

gains of $k_i = 1100$ and $k_p = 2$ for the PI controller. This also reduces the effect of measurement noise on the controlled displacement output due to feedback. The tracking performance for a 10 Hz triangular reference is shown in Fig. 10.21 (labeled "With prefilter"). It can be observed that displacement output closely follows a desired triangular reference within a wide range of 10 μm with a standard deviation of 20 nm.

10.6 Conclusion

The control issues of MEMS nanopositioning devices are introduced and discussed. A novel micro-machined silicon nanopositioner with on-chip electrothermal actuators and electrothermal sensors has been presented with nanometre resolution and low drift. The MEMS positioner was embedded in a feedback loop to realize a precise position control. Due to the nonlinear nature of the electrothermal actuator, a nonlinear inversion block was added to the feedback loop to linearize the plant. The experimental results showed that the positioner with a PI controller achieved a high degree of positioning accuracy with high robustness.

The electrothermal actuator positions the heat-sink plate, which is a suitable platform for samples located on it as a load. This one-dimensional positioner is a preliminary experimental device for the future two-dimensional positioners [45], which can be used as a nanopositioner stage for scanning probe microscopy. As the forces induced by thermal expansion of the arms are large, it is expected that the performance of the positioner would not significantly change if small probes or samples are attached to the platform. Due to the limitation of the SOIMUMPs fabrication process, it is not possible to design a two-dimensional thermal-based positioner in a single device layer. However, the thermal sensors and actuators can be placed in different layers, which can be fabricated in METALMUMPs process through MEMSCAP.

Acknowledgment This research is funded by Australian Research Council (ARC) discovery grant- DP0774287 and the Griffith University Research Infrastructure Program fund from Griffith University, Australia.

References

1. G. Binnig, H. Rohrer, The scanning tunneling microscope. Sci. Am. **253**, 50–56 (1986)
2. G. Binnig, C. Quate, C. Gerber, Atomic force microscope. Phys. Rev. Lett. **56**(9), 930–933 (1986)
3. A. Pantazi, A. Sebastian, G. Cherubini, M. Lantz, H. Pozidis, H. Rothuizen, E. Eleftheriou, Control of MEMS-based scanning-probe data-storage devices. IEEE Trans. Nanotechnol. **15**(5), 824–841 (2007)
4. A. Pantazi, M.A. Lantz, G. Cherubini, H. Pozidis, E. Eleftheriou, A servomechanism for a micro-electro-mechanical-system-based scanning-probe data storage device. Nanotechonology **15**(10), 612–621 (2004)
5. G. Binnig, M. Despont, U. Drechsler, W. Häberle, M. Lutwyche, P. Vettiger, H.J. Mamin, B.W. Chui, T.W. Kenny, Ultrahigh-density atomic force microscopy data storage with erase capability. Appl. Phys. Lett. **74**(9), 1329–1331 (1999)
6. M. Despont, J. Brugger, U. Drechsler, U. Dürig, W. Häberle, VLSI-NEMS chip for parallel AFM data storage. Sensors Actuators A Phys. **80**(2), 100–107 (2000)
7. P. Vettiger, G. Cross, M. Despont, U. Drechsler, U. Dürig, B. Gotsmann, W. Häberle, M. Lantz, H. Rothuizen, R. Stutz, G. Binnig, The 'millipede'-nanotechnology entering data storage. IEEE Trans. Nanotechnol. **1**(1), 39–55 (2002)
8. Y.K. Yong, S.S. Aphale, S.O. Moheimani, Design, identification, and control of a flexure-based XY stage for fast nanoscale positioning. IEEE Trans. Nanotechnol. **8**(1), 46–54 (2009)
9. Y.K. Yong, S.S. Aphale, S.O. Moheimani, Atomic force microscopy with a 12-electrode piezoelectric tube scanner. Rev. Sci. Instrum. **81**(3), 033701 (2010)
10. B.J. Kenton, K.K. Leang, Design and control of a three-axis serial-kinematic high-bandwidth nanopositioner. IEEE/ASME Trans. Mechatron. **17**(2), 356–369 (2012)
11. S. Schonhardt, J.G. Korvink, J. Mohr, U. Hollenbach, U. Wallrabe, Optimization of an electromagnetic comb drive actuator. Sens. Actuators A Phys. **154**(2), 212–217 (2009)
12. C.H. Kim, Y.K. Kim, Micro XY-stage using silicon on a glass substrate. J. Micromech. Microeng. **12**(2), 103–107 (2002)
13. C.K. Pang, Y. Lu, J. Chen, H. Zhu, J. Yang, J. Mou, J. Guo, B.M. Chen, T.H. Lee, Design, fabrication, sensor fusion, and control of a micro X-Y stage media platform for probe-based storage systems. Mechatronics **19**, 1158–1168 (2009)

14. N. Sarkar, R. Mansour, A CMOS-MEMS scanning probe microscope with integrated position sensors, in *1st Microsystems and Nanoelectronics Research Conference, 2008. MNRC 2008*, pp. 77–80, 15 October 2008. doi:10.1109/MNRC.2008.4683382
15. N. Sarkar, R.R. Mansour, O. Patange, K. Trainor, CMOS-MEMS atomic force microscope, in *2011 16th International Solid-State Sensors, Actuators and Microsystems Conference (TRANSDUCERS)*, pp. 2610–2613, 5–9 June 2011. doi:10.1109/TRANSDUCERS.2011.5969836
16. D. Barrettino, S. Hafizovic, T. Volden, J. Sedivy, K. Kirstein, A. Hierlemann, H. Baltes, CMOS monolithic atomic force microscope, in *2004 Symposium on VLSI Circuits. Digest of Technical Papers*, pp. 306–309, 17–19 June 2004. doi:10.1109/VLSIC.2004.1346597
17. S. Devasia, E. Eleftheriou, S.O. Moheimani, A survey of control issues in nanopositioning. IEEE Trans. Control Syst. Technol. **15**(5), 802–823 (2007)
18. A. Sebastian, D. Wiesmann, Modeling and experimental identification of silicon microheater dynamics: a systems approach. J. Microelectromech. Syst. **17**(4), 911–920 (2008)
19. Y.C. Chen, R.T. M'Closkey, T.A. Tran, B. Blaes, A control and signal processing integrated circuit for the JPL-Boeing micromachined gyroscopes. IEEE Trans. Control Syst. Technol. **13**(2), 286–300 (2005)
20. V. Todorov, V. Stavrov, J. Kreuter, Sub nm-resolution static measurement with MEMS displacement sensors. Procedia Eng. **25**, 591–594 (2011)
21. L. Gu, X. Li, H. Bao, B. Liu, Y. Wang, M. Liu, Z. Yong, B. Cheng, Single wafer processed nanopositioning XY-stage with trench-sidewall micromachining technology. J. Micromech. Microeng. **16**(7), 1349–1357 (2006)
22. C.S.B. Lee, S. Han, N.C. MacDonald, Single crystal silicon (SCS) XY-stage fabricated by DRIE and IR alignment, in *The Thirteenth Annual International Conference on Micro Electro Mechanical Systems, 2000. MEMS 2000*, pp. 28–33, 23–27 January 2000. doi:10.1109/MEMSYS.2000.838485
23. A.G. Fowler, A.N. Laskovski, A.C. Hammond, S.O. Moheimani, A 2-DoF electrostatically actuated MEMS nanopositioner for on-chip AFM. J. Microelectromech. Syst. **21**(4), 771–773 (2012)
24. J.J. Gorman, Y. Kim, N.G. Dagalakis, Control of MEMS nanopositioners with nano-scale resolution, in *Proceeding of IMECE2006*, Chicago, 2006
25. R. Hickey, D. Sameoto, T. Hubbard, M. Kujath, Time and frequency response of two-arm micromachined thermal actuators. J. Micromech. Microeng. **13**(1), 40–46 (2003)
26. Y. Sun, M.A. Greminger, D.P. Potasek, B.J. Nelson, A visually servoed MEMS manipulator, in *Experimental Robotics VIII* (Springer, Berlin/Heidelberg, 2003), pp. 255–264
27. L.L. Chu, Y.B. Gianchandani, A micromachined 2D positioner with electrothermal actuation and sub-nanometer capacitive sensing. J. Micromech. Microeng. **13**(2), 279–285 (2003)
28. Y.S. Choi, Y. Zhang, D.W. Lee, A thermal-driven silicon micro xy-stage integrated with piezoresistive sensors for nanopositioning. J. Micromech. Microeng. **22**(5), 055002 (2012)
29. L. Sun, J. Wang, W. Rong, X. Li, H. Bao, A silicon integrated micro nano-positioning XY-stage for nano-manipulation. J. Micromech. Microeng. **18**(12), 125004 (2008)
30. L. Ji, Y. Zhu, S.O. Moheimani, M.R. Yuce, A micromachined 2DOF nanopositioner with integrated capacitive displacement sensor, in *Sensors, 2010 IEEE*, Hawai, 2010
31. M.A. Lantz, G.K. Binnig, M. Despont, U. Drechsler, A micromechanical thermal displacement sensor with nanometre resolution. Nanotechnology **16**(8), 1089–1094 (2005)
32. G.K. Binnig, M. Despont, M.A. Lantz, P. Vettiger, Thermal movement sensor. International Patent WO 2004/020328 A1
33. N.B. Hubbard, M.L. Culpeper, L.L. Howell, Actuators for micropositioners and nanopositioners. Appl. Mech. Rev. Trans. ASME **59**(6), 324–334 (2006)
34. Y. Zhu, S.O. Moheimani, M.R. Yuce, Simultaneous capacitive and electrothermal position sensing in a micromachined nanopositioner. IEEE Electron Device Lett. **32**(8), 1146–1148 (2011)
35. Y. Zhu, A. Bazaei, S.O. Moheimani, M.R. Yuce, A micromachined nanopositioner with on-chip electrothermal actuation and sensing. IEEE Electron Device Lett. **31**(10), 1161–1163 (2010)

36. A. Sebastian, A. Pantazi, S.R. Moheimani, H. Pozidis, E. Eleftheriou, Achieving subnanometer precision in a MEMS-based storage device during self-servo write process. IEEE Trans. Nanotechnol. **7**(5), 586–595 (2008)
37. M.A. Lantz, H.E. Rothuizen, U. Drechsler, W. Haberle, M. Despont, A vibration resistant nanopositioner for mobile parallel-probe storage applications. J. Microelectromech. Syst. **16**(1), 130–139 (2007)
38. V. Kaajakari, *Practical MEMS* (NV: Small Gear Publishing, Las Vegas, 2009)
39. Y. Zhu, A. Bazaei, S.R. Moheimani, M.R. Yuce, Design, modeling, and control of a micromachined nanopositioner with integrated electrothermal actuation and sensing. J. Microelectromech. Syst. **20**(3), 711–719 (2011)
40. A. Bazaei, Y. Zhu, R. Moheimani, M.R. Yuce, Analysis of nonlinear phenomena in a thermal micro-actuator with a built-in thermal position sensor. IEEE Sensors J. **12**(6), 1772–1784 (2012)
41. [Online]. Available: http://www.memscap.com/en_mumps.html
42. *Software Manual – Planar Motion Analyzer Software 2.3,* Polytec. [Online]. Available: www.polytec.com
43. U. Durig, Fundamentals of micromechanical thermoelectric sensors. J. Appl. Phys. **98**, 044906 (2005)
44. C.D. Lott, T.W. McLain, J.N. Harb, L.L. Howell, Modeling the thermal behavior of a surface-micromachined linear-displacement thermomechanical microactuator. Sensors Actuators A **101**, 239–250 (2002)
45. M. Rakotondrabe, A.G. Fowler, S.R. Moheimani, Control of a novel 2-DoF MEMS nanopositioner with electrothermal actuation and sensing. IEEE Trans. Control Syst. Technol. **22**(4), 1486–1497 (2014)

Chapter 11
A Review of Nanomanipulation in Scanning Electron Microscopes

Devin K. Luu, Chaoyang Shi, and Yu Sun

Abstract Nanomanipulation under scanning electron microscope (SEM) imaging has enabled the characterization of nanomaterials and nanostructures and the prototyping/assembly of nanoscale devices. In this chapter, techniques for nanorobotic manipulation in SEMs are reviewed. Nanomanipulation platforms, nanomanipulation tools, gas injection systems, imaging techniques for automation, and applications in nanomaterial characterization and nanodevice assembly are discussed. Many micro-tools, changeable toolboxes, and automation techniques are emerging. Nanoscale laboratories built around SEMs are becoming a powerful platform to enable flexible prototyping of nanomaterial-based devices in the nanotechnology, biotechnology, and nanoelectronics sectors.

11.1 Introduction

Nanomanipulation is a technology for manipulating objects at the nanometer scale which addresses the demand to conduct fundamental investigations in areas such as material science, biology, and electronics and to improve industrial procedures to allow rapid prototyping and inspection of nanoscale devices. For example, mechanical and electrical properties of nanowires can be measured, and they can be precisely maneuvered and placed to assemble transistors. Multi-walled nanotubes can be used as linear or rotational bearings on nanoscale mechanical structures.

Nanomanipulation technology was first demonstrated in 1990, when Donald Eigler and Erhard Schweizer of IBM used a scanning tunneling microscope (STM) to position 35 xenon atoms on a metal substrate to form the letters "IBM" [1]. Although individual atoms had been moved before, this marked the first time atoms were placed deliberately with such precision. The STM used by IBM is in a class of microscopes called scanning probe microscopes, which physically move a sharp

D.K. Luu • C. Shi • Y. Sun (✉)
Advanced Micro and Nanosystems Lab, University of Toronto, Toronto, ON, Canada
e-mail: sun@mie.utoronto.ca

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_11

tip over a surface to detect a surface profile. Many studies of nanomanipulation have been performed using another type of scanning probe microscopes, called atomic force microscopes (AFM) [2–4]. The advantage of the AFM is that it is not only capable of atomic scale resolution, but it is also much more compact. However, the disadvantage of using scanning probe microscopes alone is their limited motion range and slow image feedback. It can take several minutes to acquire an image. Even today, performing simultaneous manipulation and imaging using AFM remains a challenge.

Two other microscope technologies used to image manipulate materials at the nanometer level are transmission electron microscopes (TEM) and scanning electron microscopes (SEM). The TEM is capable of real-time, sub-nanometer resolution imaging, allowing high precision characterization and measurements [5]. However, the specimen chamber and scanning/observation workspace are small, demanding special sample preparation techniques [6]. The SEM provides real-time imaging with up to 1 nm resolution and boasts a larger specimen chamber and scanning area, allowing integration and development of robotic manipulation systems. These advantages allow simultaneous high resolution imaging, manipulation, and direct interactions with samples [7]. A comparison of high resolution imaging technologies is shown in Table 11.1.

Nanomanipulation platforms with multiple degrees of freedom piezoelectric actuator-based manipulators capable of manipulating nanoscale objects in nanometer resolution have been built specifically inside SEM to perform tasks ranging from electrical probing [8] to cell nanosurgery [9]. Electron beam induced deposition (EBID) systems can be combined with nanomanipulation platforms to perform pick-and-place and assembly tasks [10] or to directly deposit nanostructures. AFM systems have been integrated into SEM for simultaneous imaging and nanomanipulation, combining the advantages of both instruments to allow electrical and mechanical characterization of materials, SEM-guided topography imaging as well as crystallography, and chemistry and surface morphology analysis [11]. Recent progress also involves the integration of AFMs and focused ion beam (FIB) systems to form a nanoscale laboratory capable of simultaneous imaging, fabrication, and nanomanipulation [12]. This chapter reviews existing SEM nanomanipulation techniques.

Table 11.1 A comparison of high resolution imaging technologies

	Typical imaging area	Resolution	Image acquisition speed
AFM	100×100 μm	Lateral: depends on tip geometry Vertical: 0.1 nm	Minutes
TEM	3 mm diameter ring	0.1 nm in-plane	Milliseconds
SEM	70×50 mm	1 nm in-plane	Milliseconds

11.2 Scanning Electron Microscopy Imaging for Automation

The SEM is capable of providing 1 nm resolution visual feedback at up to 25 Hz with a depth of field on the order of micrometers. Not only are these properties useful for teleoperation of nanomanipulation systems, but they also make the SEM a promising candidate as a vision sensor for automation. However, many nanomanipulation systems still operate in open loop without any feedback mechanisms. Further developments in SEM imaging techniques can enable efficient and robust automation in SEM nanomanipulation. This section introduces basic SEM imaging principles and summarizes the methods for imaging de-noising, drift compensation, depth sensing for automation, and 3D sample reconstruction.

11.2.1 SEM Imaging Basics

The electron's short wavelength allows it to probe at scales much smaller than optical microscopes (OM). Where the wavelength of visible light limits OM resolution to about 200 nm, the resolution of SEM imaging can be as fine as 1 nm. The depth of field in the SEM is also much higher, as shown in Fig. 11.1 [13]. In an SEM, an electron gun generates a beam of electrons towards the sample. The condenser and objective lenses focus the beam to a point on the sample. A beam deflector controls the beam to scan across the sample. Electrons hitting the surface of the sample cause secondary or backscatter electrons to be emitted and then detected by an electron detector, which converts the incoming electrons to signals, sending them to a video screen as shown in Fig. 11.2.

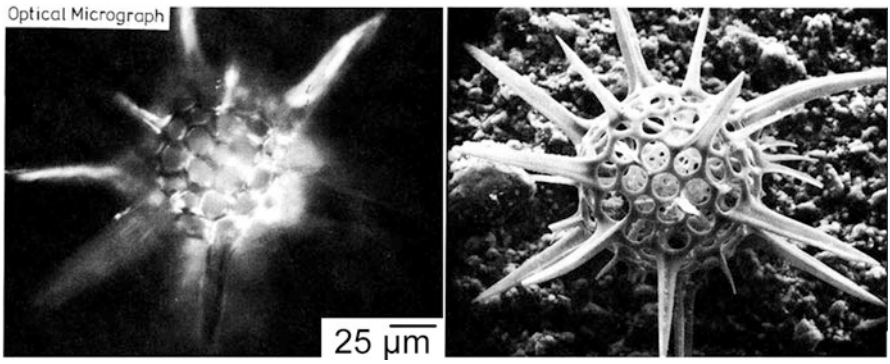


Fig. 11.1 The OM image, showing only a small cross section of the specimen, has a smaller depth of field than the SEM image, which shows the entire specimen and the background with clarity [13]

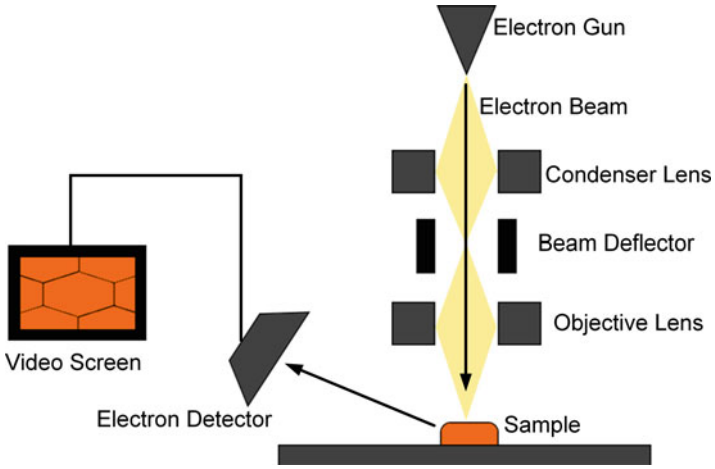


Fig. 11.2 Operation of the SEM. An electron gun fires an electron beam, which is focused by two lenses, and is maneuvered by the beam deflector to scan the sample. Electrons are emitted from the sample and detected by an electron detector, sending a signal to a video screen

Another important aspect of the SEM is its vacuum environment because air in the SEM chamber can interfere with electron paths. SEMs typically use high vacuum in the imaging/specimen chamber and use a specimen exchange port to preserve the vacuum within the SEM specimen chamber when a sample is brought into or taken out of the specimen chamber. Because the working principle is based on detecting electrons emitted from the sample, it is important to control specimen charging. Non-conductive specimens are typically sputter coated with gold to ensure no charge is built up.

11.2.2 Imaging Modes

The two main imaging modes are secondary electron (SE) mode, showing topographical information, and backscattered electron (BSE) mode, showing composition information [14]. SEs are produced when an incoming electron with sufficient energy collides with the sample to induce emission of electrons. Sharp edges on the sample emit more SEs while flatter surfaces emit fewer. BSEs are produced when an electron is emitted towards a sample, and the emitted electron's path is altered by the charge of the nuclei in the sample's atoms, similar to how the path of a comet is altered by the gravity of a star. Atoms with higher atomic numbers cause more BSEs. The difference between SE and BSE images is depicted in Fig. 11.3.

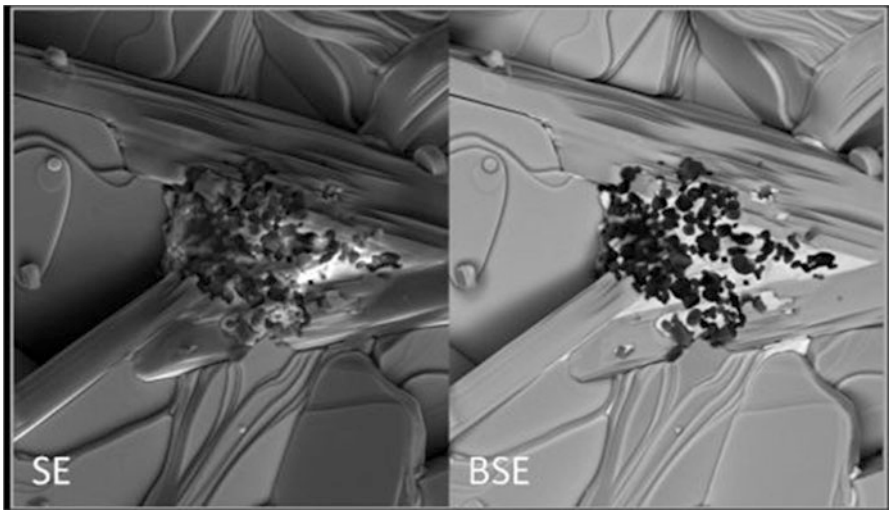


Fig. 11.3 The SE image on the *left* shows topographical information. The BSE image on the *right* shows composition of the sample, represented as different colors in the image [14]

Table 11.2 Parameters affecting SEM imaging

	Low	High
Acceleration voltage	• Wider spot size • Lower sample penetration	• Narrower spot size • Higher sample penetration • Risk of sample damage
Aperture size	• Narrower spot size • Lower signal-to-noise ratio	• Wider spot size • Higher signal-to-noise ratio
Beam current	• Fewer electrons hit sample • Lower signal-to-noise ratio	• More electrons hit sample • Higher signal-to-noise ratio

11.2.3 Parameters Affecting SEM Imaging

The three main parameters affecting imaging quality are acceleration voltage, aperture size, and beam current. These settings affect both the spot size of the electron beam and the number of electrons striking the sample. There is no all-around best setting, and the operator must tune these parameters in order to obtain the best possible image for each specific circumstance. A summary of the effects of changing the above parameters is summarized in Table 11.2.

In general, a smaller spot size gives higher resolution, and more electrons striking the sample gives a higher signal-to-noise ratio. However, when more electrons strike the sample, it becomes difficult to focus the electron beam, which can sacrifice imaging resolution. Although acceleration voltage can be increased to narrow the spot size, increasing it too much can increase the risk of sample damage, and too much sample penetration can introduce artifacts into the image via the

emission of secondary electrons from inside the sample. Therefore, there is always a trade-off between resolution and noise in SEM imaging. Accelerating voltage and beam current must also be controlled to limit charging of the sample when low conductivity samples are imaged in order to limit image drift.

11.2.4 In-plane Imaging

To use SEM images as feedback for visual servoing (i.e., vision-based control), the image must be stable and clear. There are two main difficulties preventing accurately locating in-plane features: noise and drift. Noise in SEM images obscures details in images and is caused by random factors in the ambient environment. Noise is especially prevalent when imaging delicate samples when the accelerating voltage and beam current cannot be too high so there is a low signal-to-noise ratio. Image drift occurs when the entire image shifts, which is caused either by the specimen physically shifting or the electron beam shifting. Image drift can also occur if the specimen does not conduct charge away from its surface sufficiently quickly, or if there are interfering magnetic fields. Low acceleration voltages at high magnification and short irradiation times can also cause SEM image drift.

Methods for de-noising images such as Gaussian smoothing [15], neighborhood filter [16], and anisotropic filter [17] often obscure relevant image details. Total variation minimization [18, 19] and non-local (NL) means de-noising can smooth out noise and are also effective in preserving sharp edges and details [20, 21]. The result of the NL-means de-noising algorithm applied on an SEM image is shown in Fig. 11.4. To compensate for drift, image correlation is used to track predetermined

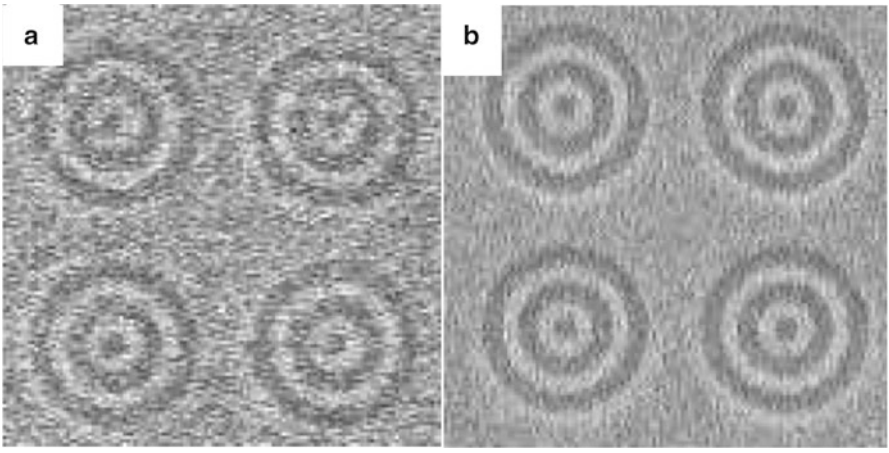


Fig. 11.4 The NL-means de-noising algorithm is applied on the SEM image. (**a**) Unprocessed image. (**b**) Processed image. Adapted from [22]

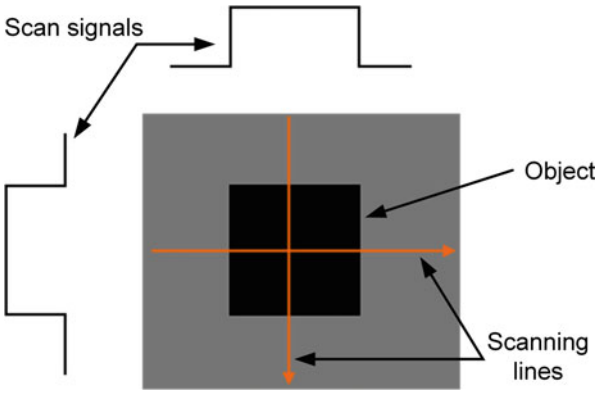


Fig. 11.5 Line scanning technique to speed up position feedback acquisition

stationary points in an image [22]. If the selected features move, then the image is shifted back using correlation-based template matching. Gong et al. [22] used the NL-means method and image correlation to condition SEM visual feedback for an automated nanoprobing task.

11.2.5 In-plane Imaging Speed

Although the frame rate of SEM imaging can be as high as 25 Hz, automation and controls applications can benefit from a faster feedback rate. If the geometry of the object is known, a line scan approach can be used instead of acquiring the entire image to speed up automated tasks within the SEM, as shown in Fig. 11.5 [23]. Using this approach, Tiemerding et al. tracked moving rectangular features and achieved a 1 kHz update rate [23].

11.2.6 Depth Sensing in SEM

Depth sensing, which is difficult to achieve in SEM due to its large depth of field, is important in microscopy applications, especially when automation is required in 3D because nanomanipulation tools are delicate and easy to break. Hence, it is imperative that the vertical position of tools is precisely known. Opening the SEM chamber to replace a broken tool also incurs lengthy pumping in order to restore vacuum in the chamber. Because the method of depth sensing used in OMs in which a tool tip is brought into focus on a reference plane of known depth is ineffective in SEM due to its large depth of field, two main ways to detect depth have been developed.

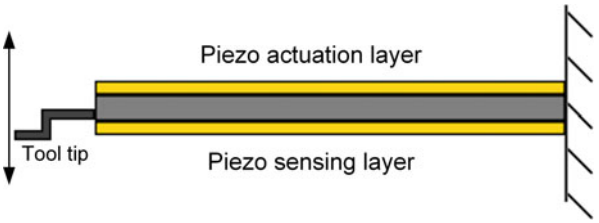


Fig. 11.6 Vibrating bimorph actuator for contact detection. When contact is established, the amplitude of vibration drops

11.2.6.1 Contact Detection

Contact sensor-based depth sensing uses a probe or force sensor to touch a reference plane of known depth and can be as accurate as 20 nm [24]. The AFM can be used to detect contact because it inherently outputs force or height information, such as via piezoresistive signals [25]. Eichhorn et al. [26] implemented a touch sensor using a vibrating piezoelectric bimorph actuator with an integrated sensor as shown in Fig. 11.6. The vibration amplitude distinctly drops upon contact and a sensor signal drop is detected. However, because the tool tip is connected to the bimorph actuator, there are potentially unwanted vibrations.

Ru and To [24] implemented a contact detection method by detecting tool tip sliding. A probe is moved slowly towards a surface and upon contact, sliding motion is detected from image processing. Electrical connections are also attached to the probe, which can be used to electrically verify touch. Technologies such as piezoresistive and capacitive sensors could provide high resolutions and reliability; however, they are difficult to integrate with nanomanipulation tools [24].

11.2.6.2 Vision-Based Depth Sensing

Vision-based depth sensing utilizes various strategies to determine the depth of the tool tip from images alone. Kasaya et al. [27] integrated an optical microscope to the wall of the SEM chamber to detect the depth of the tool tip. Cvetanovic et al. used a similar approach, but instead of using an optical microscope, an infrared camera and a scale were fixed to the wall of the chamber [28] as shown in Fig. 11.7.

Stereo image information can be used to determine depth of the tool tip by tilting the sample stage [29, 30], or by tilting the electron beam [31]. Stereo imaging in SEM requires careful mechanical design. Tilting the electron beam requires additional instrumentation near the beam emitter for beam control. Additionally, correlating two or multiple images must be well handled due to SEM image noise. Differently, Eichhorn et al. developed a shadow-based depth detection process [32, 33]. However, occlusions and poor shadow contrast from image background cause difficulties in image segmentation.

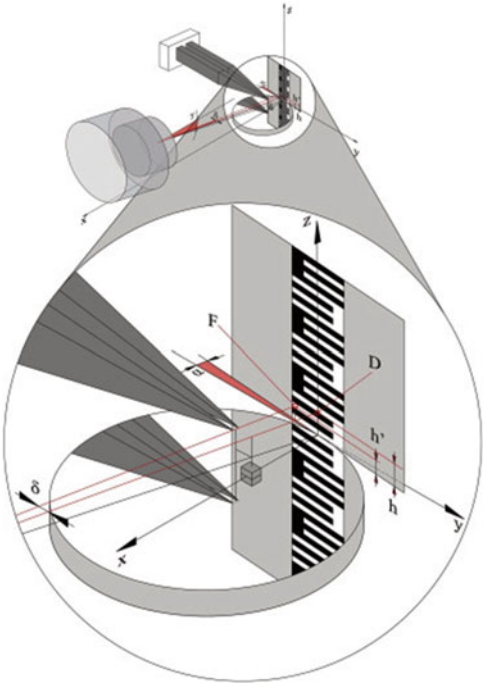


Fig. 11.7 A scale is mounted on the SEM chamber wall, and an IR camera is used to detect the tool tip position relative to the scale [28]

11.2.7 Ultramicrotomy for 3D Image Reconstruction

Ultramicrotomy is a relatively new application in SEM nanomanipulation and can be used to create 3D reconstructions of blocks of tissue. It is an important process in histology, the study of the anatomy of cells and other microscopic structures in plants and animals. In traditional ultramicrotomy, a highly skilled technician painstakingly takes thin slices of tissue on the order of 70 nm in thickness that are then collected on a TEM grid from a small container of water attached to the knife. The technician risks either losing, wrinkling, or mixing the order of the slices.

To address these difficulties, Denk and Horstmann developed a technique called Serial Block Face Scanning Electron Microscopy (SBFSEM) to automatically slice and collect ultramicrotome images, and to reconstruct tissue in 3D from images [34]. One major advantage of SBFSEM besides automation is that the thin slices are discarded and the freshly exposed block face is imaged, eliminating the need to collect the delicate sections, as shown in Fig. 11.8. After the cutting process, computer vision algorithms are used to identify the outlines of structures in each image, and then they are used to reconstruct the tissue in 3D, as shown in Fig. 11.9. One exciting research direction following this development is the mapping of neurons in brains. Researchers have reconstructed 3D blocks of tissue from SBFSEM data [35] and are now striving for full brain maps [36].

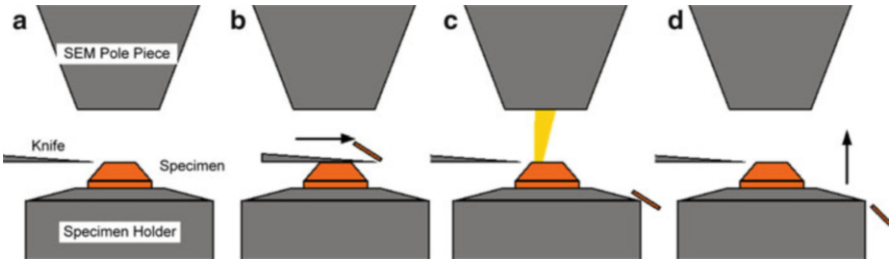


Fig. 11.8 SBFSEM process. (**a**) Starting position. (**b**) The knife moves in to take a slice from the specimen. (**c**) The SEM images the exposed block face of the specimen. (**d**) The specimen advances to prepare for another slice and the process repeats

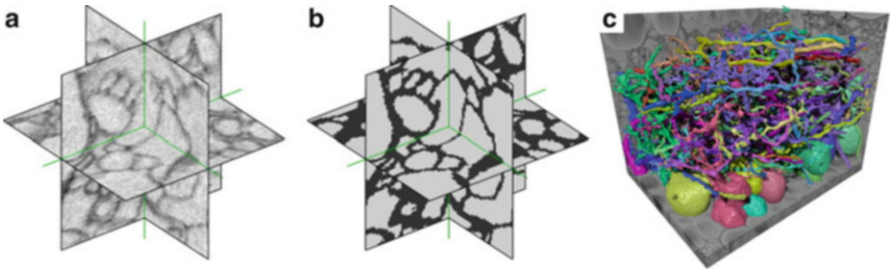


Fig. 11.9 Reconstruction of tissue in 3D. (**a**) Cross sectional images are obtained and stacked together. (**b**) Image segmentation is done by software. (**c**) From the results of image segmentation, tissue is reconstructed in 3D [35]

11.3 Nanomanipulation Systems

Nanomanipulation systems are platforms integrating one or more nanopositioners that are used in SEM to perform nanomanipulation tasks. Positioners are typically driven by piezoceramic actuators, which are both vacuum compatible and antimagnetic. Nanotools are attached to nanopositioners as "hands" to perform various tasks. Nanomanipulation systems often incorporate a combination of coarse and fine positioners to achieve greater range, speed, and accuracy. Coarse positioners are often based on the slip-stick principle are used to achieve range and speed, as shown in Fig. 11.10a–c. By expanding or contracting the piezoelectric stack abruptly, the friction interface between the pusher and moving part slips. Then, by actuating the piezoelectric stack slowly in the opposite direction, the moving part "sticks" and pure displacements are produced. However, stick-slip actuation has relatively low resolution (50 nm), and the driving principle causes vibrations. When the tool tip is approximately in position, fine positioners, based on flexure mechanisms, are used to accurately position the tool tip as shown in Fig. 11.10d. A slip-stick mechanism can also be used instead of a flexure mechanism as a fine positioner if the piezoelectric actuator is only slowly contracted and expanded, but the motion range becomes much smaller.

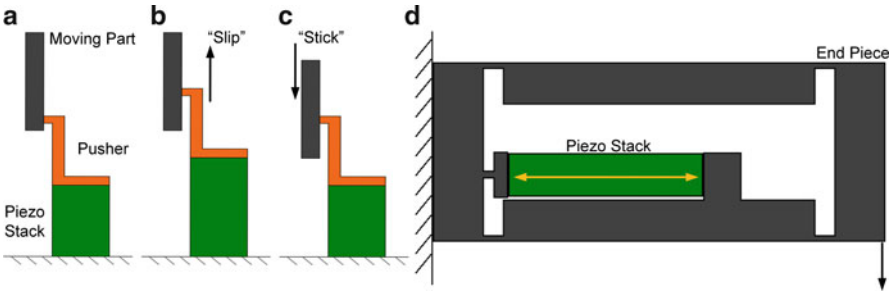


Fig. 11.10 (**a**–**c**) The slip stick mechanism works by quickly expanding the piezo stack, then slowly retracting it. The friction interface between the pusher and the moving part will slip during the quick expansion, and stick during the slow retraction. (**d**) The flexure mechanism works as a lever

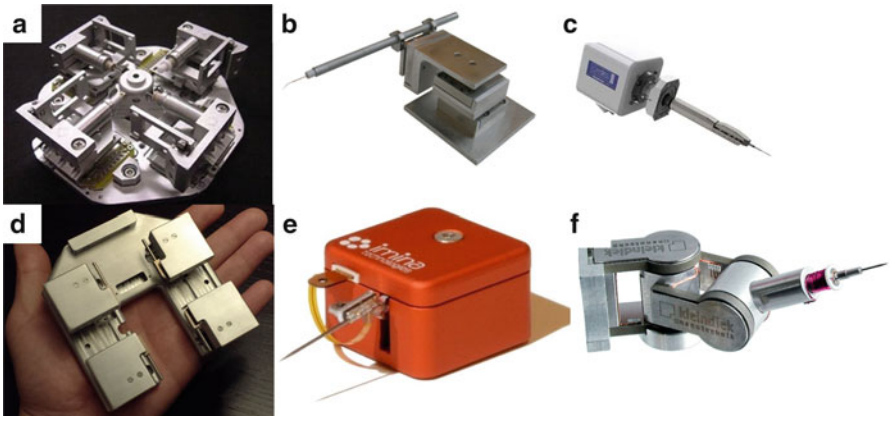


Fig. 11.11 Examples of commercial nanomanipulation systems: (**a**) Zyvex S100. (**b**) Xidex NanoBot. (**c**) Oxford Instruments OmniProbe 400. (**d**) Toronto Nanoinstrumentation LifeForce-2000 (**e**) Imina miBot BT-11 (**f**) Kleindiek Nanotechnik MM3A

Using the piezoelectric actuator-based positioners, a number of commercial nanomanipulation systems have been developed, each with different capabilities such as joystick or game controller teleoperation, closed-loop controlled, in-situ tool-tip changing, or specimen exchange port compatibility [37–42]. Some of the commercial nanomanipulation systems are shown in Fig. 11.11 and Table 11.3.

11.3.1 Hybrid AFM/FIB/SEM Systems

A number of integrated systems have emerged to form powerful nanoscale laboratories. SEM and AFM are complementary techniques for performing material surface investigations [43]. The integration of AFMs inside SEMs combines advantages of

Table 11.3 Summary of commercial nanomanipulation systems

	Speed	Motion range	Axes	Coarse resolution	Fine resolution
Zyvex S100	3 mm/s	12 mm XYZ	12	100 nm open loop	5 nm open loop
Xidex NanoBot NX-2000	1 mm/s	12 mm XYZ	6	50 nm open loop	1 nm open loop
Oxford Instruments OmniProbe 400	N/A	4 mm XYZ	4 (1 rotation)	10 nm closed loop	N/A
TNI LF-2000	>10 mm/s	10 mm XY 5 mm Z	12	<100 nm open loop	0.1 nm open loop 1 nm closed loop
Imina miBot BT-11	2.5 mm/s	Unlimited XY 10 mm Z	2 translation 2 rotation	40 nm open loop	0.5 nm open loop
Kleindiek MM3A	2 mm/s telescoping 10 mm/s rotation	12 mm telescoping 240° rotation	1 translation 2 rotation	5 nm open loop rotation 0.5 nm open loop translation	N/A

both instruments to perform high resolution analysis in real time, revealing details that may be missed by AFM or SEM alone. FIB is a standard micromachining tool using FIBs for cutting or depositing materials. Several hybrid AFM/SEM and AFM/FIB/SEM systems have been developed [3, 4, 12], and commercial systems are also available [44–48].

Conventional AFMs based on laser beam deflection are not typically integrated inside a SEM due to the space and optical path constraints. A commercial hybrid AFM–SEM combination system by DME–SPM was designed with a modified laser path inside SEM [46], shown in Fig. 11.12a. However, laser alignment in such a configuration is complicated, and the laser power must be kept low to limit thermal generation in the vacuum chamber. Attocube Systems AG used a fiber-optic configuration to construct an in-situ AFM for SEM (attoAFM/SEM) with a laser interferometer [48], in which the laser source and detector are placed outside the vacuum chamber to avoid the heating problem, shown in Fig. 11.12b. The optical fiber, AFM cantilever, and sample are tilted by 45°, enabling full SEM visual access to both AFM tip and sample.

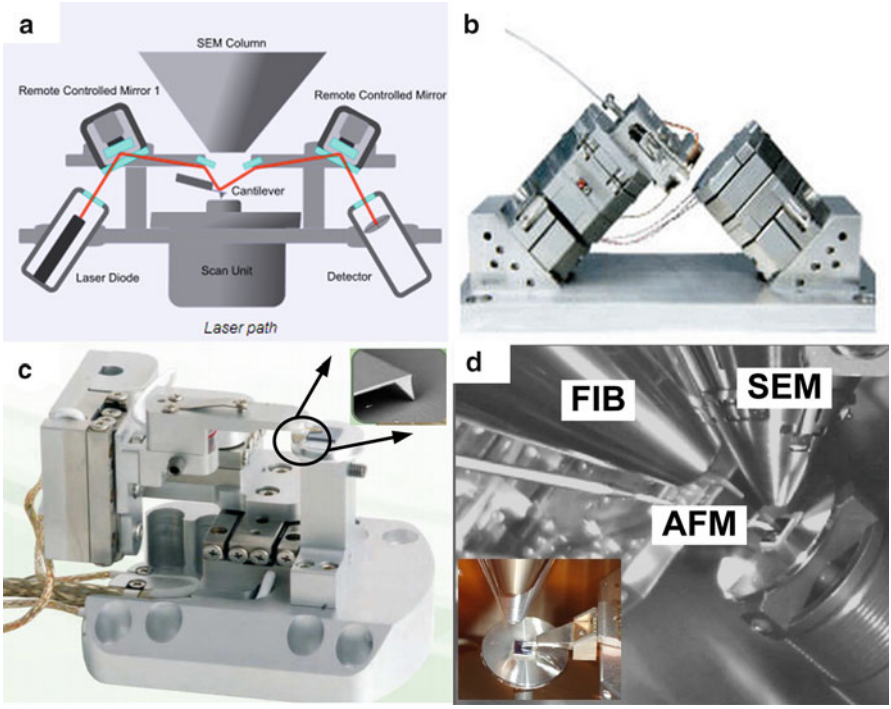


Fig. 11.12 Hybrid AFM/SEM and AFM/SEM/FIB systems. (**a**) AFM–SEM based on laser beam deflection with modified laser path system by DME-SPM. (**b**) AFM–SEM with a fiber-optic configuration by Attocube Systems AG. (**c**) AFM using tuning fork with a QPlus or Akiyama probe by Trioptics. (**d**) 3TB4000 AFM/FIB/SEM from Nanonics Imaging Ltd

In addition to optical-based methods, AFMs based on self-sensing cantilevers and tuning forks have also been used for imaging in either contact mode or tapping mode. Fatikow et al. developed an AFM subsystem using self-sensing piezoresistive cantilevers integrated inside an SEM to perform both in-situ scanning and manipulation [3], providing force feedback without using laser measurement. Dynamic working mode AFMs using tuning fork with a QPlus or Akiyama probe configuration were also integrated inside SEM to realize surface topography with a large scanning area of 500 $\mu m \times 500$ μm [47], as shown in Fig. 11.12c. The 3TB4000 from Nanonics Imaging Ltd is an integration of AFM based on a tailor-made Q-Plus tuning fork, SEM, and FIB, shown in Fig. 11.12d. This instrument allows a large field of view imaging from SEM, followed by high resolution AFM imaging and 3D fabrication with ion beam milling [44]. Fatikow et al. [12] designed a hybrid AFM/FIB/SEM system to perform similar functions using piezoresistive cantilevers in contact mode.

11.3.2 Nanotools as End-Effectors for Nanomanipulation

There are many nanotools available for nanomanipulation within SEM. The most common tasks are nanoassembly, nanomaterial and cell mechanical and electrical characterization, and cell nanosurgery.

AFM tips can be used as-is or modified to adapt to different applications in SEM nanomanipulation. Due to commercial availability and existing hardware for precise position control, hybrid AFM–SEM systems are used in nanomanipulation [3, 4, 12]. The AFM cantilever also inherently provides force feedback, which is useful for detecting contact with specimens [3]. Combined with the fast visual feedback from the SEM, AFM is suitable for both automation and teleoperation. A brief overview of different AFM tips is given in Table 11.4. The modified AFM tips are typically manufactured by FIB milling or deposition.

Non-contact tuning fork-based AFMs using an oscillating quartz tuning fork from an ordinary wristwatch have also been used for imaging [55]. In non-contact mode, the quartz tuning fork with a small probe attached to it is brought close to the sample while it is resonating, as illustrated in Fig. 11.13a. When the tuning

Table 11.4 Different AFM nanotools and their applications

Tip	Usage examples
Original tip	• DNA contact pushing [49] • Carbon nanotube contact pushing [50, 51] • Nanoparticle contact pushing [2, 3]
Flat tip	• Cell stiffness characterization [52]
Needle tip	• Cell stiffness characterization [52]
Buffering beam	• Cell stiffness characterization [53] • Single-cell cutting [54]

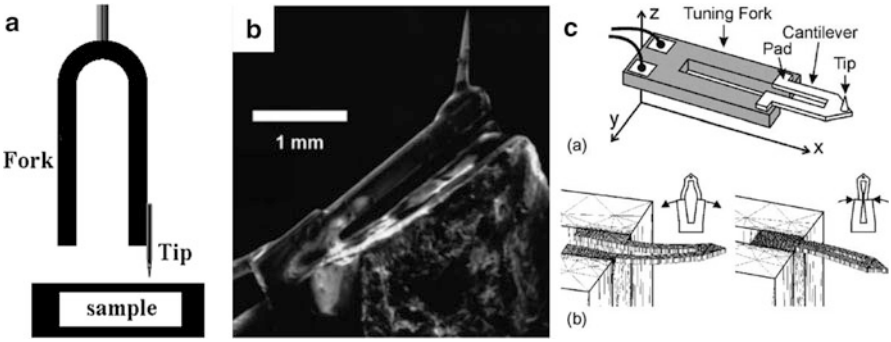


Fig. 11.13 Non-contact AFM tools: (**a**) Tuning fork [55]. (**b**) Q-Plus probe [56]. (**c**) Tapping mode AFM tool [57]

fork approaches close enough to the sample, it experiences intermolecular forces, causing a decrease in vibration amplitude and increase in frequency that can be detected as a change in a voltage signal. Another non-contact AFM called the Q-Plus probe fixes one side of the tuning fork such that only one prong vibrates as shown in Fig. 11.13b [56]. The advantage of the Q-Plus probe is that because there is only one vibrating prong, the dynamics of the system can be easier to interpret.

Tapping mode AFMs operate with intermittent contacts. The Akiyama probe, a type of tapping mode AFM tool, is a cantilever bonded to a tuning fork which moves up and down in a tapping motion as the tuning fork vibrates, as shown in Fig. 11.13c [57]. Noncontact and tapping mode AFM are more suitable for nondestructive measurement to perform characterization on delicate samples because the probe does not physically touch the sample as severely or often as in the contact mode. The cantilever on the Akiyama probe can also be modified to adjust stiffness, allowing delicate biological samples to be probed.

A tungsten probe is a piece of metal with a nanometer tip, originally used for making electrical measurements [58, 59]. At nanometer scales, intermolecular and electrostatic forces are more dominant [60], making probes a versatile tool for many nanomanipulation tasks such as pick and place for nanoassembly, especially for manipulating carbon nanotubes [61] and nanowires [10]. Additionally, microelectromechanical systems (MEMS) micro/nano-grippers have also been developed and used in SEM nanomanipulation for gripping nanoparticles [62].

11.3.3 Nanotool Changer

Nanomanipulation can require the use of different tool tips. Broken tools also need to be replaced before manipulation can proceed. Thus, a toolbox array with different tip morphologies was built, shown in Fig. 11.14a. The tools, called Nanobits, are fabricated out of Si_3N_4 or SiO_2 by electron beam lithography and

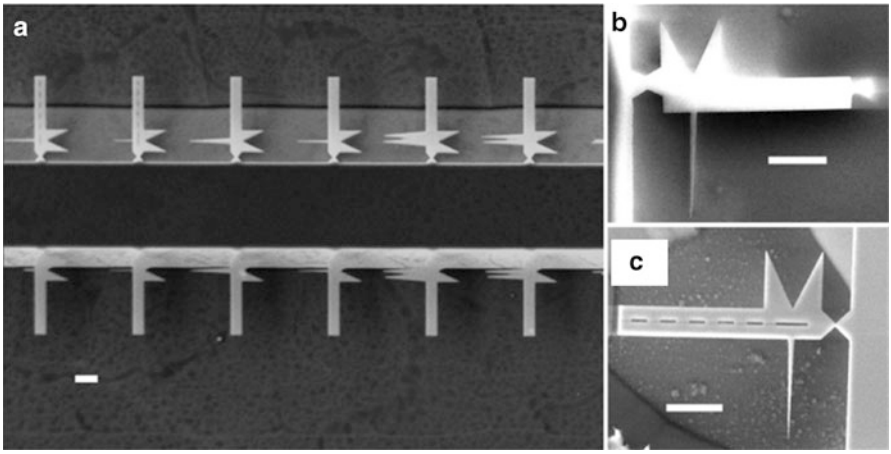


Fig. 11.14 Changeable toolbox inside SEM and ESEM: (**a**) A library of free-standing nitride Nanobits with different tip morphologies. (**b**) Silicon dioxide Nanobit. (**c**) Silicon nitride nanobit [64]

silicon processes and are 2–4 μm long and 120–150 [63, 64]. Two typical tools are shown in Fig. 11.14b, c. In operation, a microgripper detaches these tools from the array and assembles them to an AFM probe or other end effectors. Another nanotool exchange system was proposed in ESEM for biological applications [65]. Low-melting temperature metal was used to attach and hold different tools for cell characterization and surgery. The Oxford OmniProbe 400 is a commercial nanomanipulator which is also capable of in-situ tool exchange and probe tip sharpening for repairing or replacing broken tools without opening the SEM chamber [39]. Using an in-situ tool changer can reduce the amount of time required to change a tool tip from hours to minutes.

11.4 Gas Injection for Nanomanipulation

Chemical reactions via gases injected into the SEM chamber can be used to perform useful assembly tasks. Gases can be used to deposit structures using the EBID process. Small amounts of gas can also be introduced near a non-conductive sample to whisk away built-up charges.

11.4.1 Electron Beam Induced Deposition

In EBID, precursor particles adsorbed on a substrate are dissociated by a focused electron beam, thereby depositing material on the substrate. The resulting deposited structures are nanometers in scale. EBID's ability to deposit metals, dielectrics, and semiconductor materials makes this technology suitable for applications such as nanoassembly to join materials [10], circuit, or photomask alterations or repairs [66], and sensor applications. Visual inspection can be performed directly after the EBID process.

An EBID system consists of two key parts: an electron beam source and a nozzle. A substrate surface is placed inside a chamber with the nozzle and the electron beam. The nozzle, a fine capillary connected to a gas source, injects precursor molecules such as WO_3 onto the substrate where they are adsorbed. The electron beam source emits a directed beam at the locations where deposition is desired, dissociating the precursor molecules. The deposit (e.g., tungsten) remains on the surface of the substrate while the dissociated products (e.g., oxygen) are desorbed. A graphical illustration of the EBID process is shown in Fig. 11.15.

The resolution of EBID in SEM can be as fine as 3 nm [67]. Although the electron beam can be focused with a resolution finer than 1 nm, the action of dissociation is actually done by secondary electrons diffracted from the surface of the sample. The minimum feature size is therefore a function of the electron beam and the spread of diffracted secondary electrons. The reason for this phenomenon is the so-called cross-section for the precursor molecules. The cross-section is a measure of how likely particles are to interact, and secondary electrons possess the optimal energy to maximize cross-section. For example, the electron beam from an SEM has an energy of thousands of eV, whereas the secondary electrons only have energies of tens of eV.

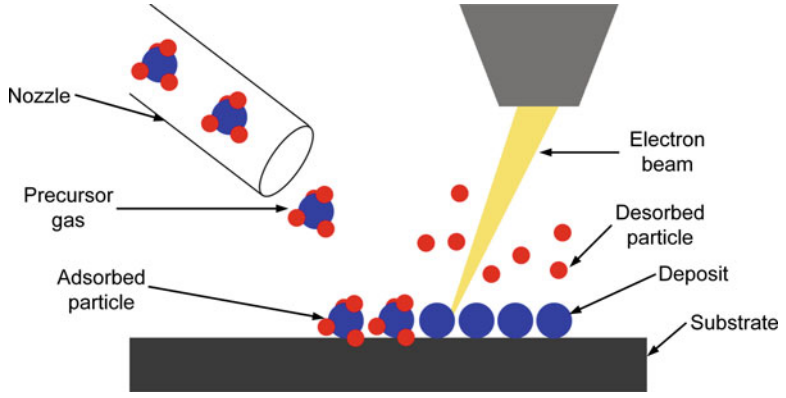


Fig. 11.15 Gas source injects precursor molecules from a small nozzle (e.g., 0.5 mm diameter). Molecules are adsorbed on substrate. Directed electron beam dissociates precursor molecule. Particles are deposited on surface of substrate. Dissociated products are desorbed

Table 11.5 Parameters that affect EBID

Parameter	Effect
Beam energy (E)	Use high energy for high resolution [69]
Beam current (I)	Use high current for high deposition rate [70] according to rate equation model
Dwell time (t_D)	A long dwell time will cause undesired particles that have not yet been desorbed to be deposited [69]
Replenishment time (t_r)	Higher replenishment rate will increase the partial pressure of the precursor, increasing resolution and deposition rate [71]

Table 11.6 Common EBID precursor materials

Material	Remarks
Metal-halogen complexes (WF_6)	Highly toxic and corrosive
Tungsten hexacarbonyl ($W(CO)_6$)	For tungsten deposition
Metal carbonyls ($Me(CO)_x$)	Metal deposition, carbon contamination
Trimethylcyclopentadienyl platinum ($CpPtMe_3$)	For platinum deposition

EBID can also deposit microstructures such as nanogranular structures, which makes EBID suitable for developing strain sensors and magnetic sensors [68]. Other possible deposit microstructures include amorphous and polycrystalline structures, each of which exhibits different electric, magnetic, and mechanical properties [68]. The resulting structure depends on process control and the precursor molecule used. Organometallic precursors with tendency to form carbides usually result in amorphous structures [68]. Metallic structures unlikely to form carbides usually result in nanogranular structures, composed of metallic particles in a dielectric matrix. For certain precursors, a polycrystalline deposit microstructure may be obtained by allowing only small carbon impurity depositions [68].

There are several parameters that affect the quality or speed of deposition. The beam energy, current, dwell time, and replenishment time affect deposition rates and resolution, as shown in Table 11.5.

However, the most important parameter in the EBID process is the selection of the precursor. A good precursor material must produce a solid deposit after the dissociation process, and the deposit must exhibit the required physical properties. Common precursor gases and their applications are shown in Table 11.6.

There are commercial EBID systems for SEM, which consist of one or more nozzles connected through the SEM feedthrough port to reservoirs of precursor gases. A motion stage moves the EBID nozzle to points of interest. Examples of commercial gas injection systems are the Xidex PGIS [72], Oxford Instruments OmniGIS II [73], and the Orsay Physics GIS-5 [74], which can all supply at least three precursor gases. Some commercial gas injection systems are shown in Fig. 11.16.

Nanowire growth via field emission is a nanoscale fabrication technique based on the decomposition of precursor gas molecules absorbed on a substrate by field

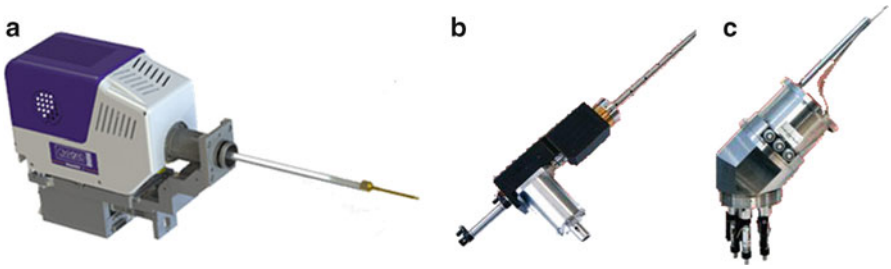


Fig. 11.16 Examples of EBID gas injection systems: (**a**) Oxford instruments OmniGIS. (**b**) Orsay physics MonoGIS. (**c**) Orsay Physics GIS-5

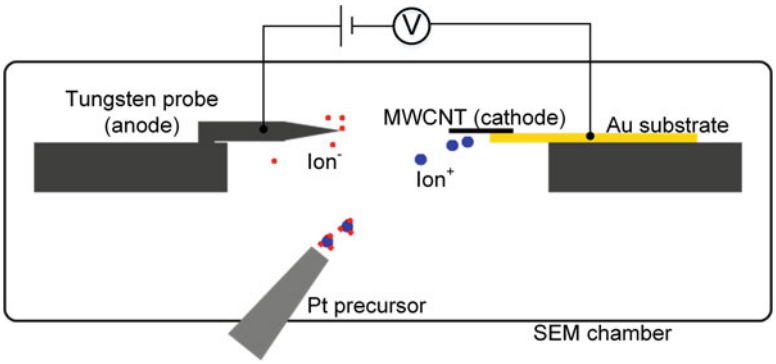


Fig. 11.17 High-purity platinum nanowires are grown on a cathode from gas injection. Adapted from [76]

emission current. The precursor molecules can decompose into positive ions and negative ions, which are then attracted and deposited to cathodes and anodes by electrostatic force. This technique produces high purity and hybrid metallic nanowire growth [75, 76], and various types of nanowires can be grown with the introduction of different precursors.

A high-purity platinum nanowire was fabricated via exposure inside a field emission electron microscope using this technique [75]. A multi-walled CNT (MWCNT) worked as a cathode, providing an emitter for a low working field emission voltage, and a tungsten probe was employed as a field emission anode. After introducing trimethylcyclopentadienyl platinum ($CpPtMe_3$), the platinum was deposited on the emitter to form a platinum nanowire, shown in Fig. 11.17. In a similar procedure, a metallic sensor consisting of a platinum and tungsten dual nanowire electrode for intracellular pH detection was fabricated with the introduction of $CpPtMe_3$ and tungsten hexacarbonyl ($W(CO)_6$) respectively, on two different MWCNT emitters [77, 78]. The tungsten nanowire is used as a work electrode and the platinum nanowire is used as a counter electrode. The performance of the pH sensor was tested inside acid and alkali buffer solutions [78].

11.5 Nanomanipulation Application Examples

The nanometer positioning resolution of SEM nanomanipulation systems has proven useful for a number of applications such as mechanical and electrical properties characterization, nanoassembly, and cell nanosurgery.

11.5.1 Nanomaterials Characterization

Characterization of nanoscale materials such as silicon nanowires, carbon nanotubes (CNT), and graphene sheets are important for investigating their properties and exploring applications of these materials. Various nanoscale mechanical and electrical characterization experiments have been performed with nanorobotic manipulation inside SEM.

Fukuda et al. performed an in-situ mechanical property characterization experiment to determine the Young's modulus of an individual MWNT, as shown in Fig. 11.18a. A bundle of MWNTs was soldered on an AFM cantilever by EBID and then buckled on a substrate by pushing the cantilever down [61]. In a similar procedure, Zhu et al. demonstrated in-situ tensile testing for mechanical characterization of silver nanowires. After two ends of the nanowire were fixed to

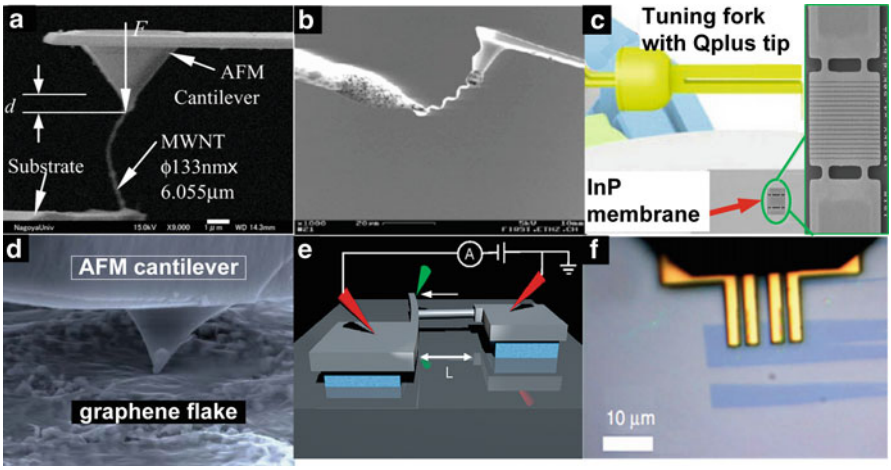


Fig. 11.18 Manipulation of nanoscale materials for property characterization. (**a**) A bundle of MWNTs was buckled for Young's modulus determination. Adapted from [61] (**b**) A nanospring was stretched between a probe and an AFM cantilever. [80] (**c**) Tuning fork AFM to characterize InP membrane [81] (**d**) Nanoindentation test for mechanical characterization of a graphene flake. Adapted from [84] (**e**) Schematic illustration for piezoresistivity characterization of a suspended Si nanowire with tensile testing. Adapted from [86] (**f**) Electrical characterization of graphene flakes. Adapted from [87]

an AFM cantilever and a tungsten probe by carbonaceous material deposition by EBID, tensile testing was performed by moving the cantilever and tungsten probe apart. A range of mechanical properties was measured, including Young's modulus, yield strength, and ultimate tensile strength [79]. An InGaAs/GaAs nanospring was stretched for tensile tests using a similar method for mechanical characterization [80], shown in Fig. 11.18b.

Figure 11.18c shows an example application for the characterization of 2D nanomaterials. Abrahamians et al. [81] performed a nondestructive measurement method using a tuning-fork-based dynamic force sensor inside a SEM to determine the stiffness of a batch of suspended InP micromembranes. The tuning fork with a QPlus configuration was controlled with frequency modulation mode, and the frequency shift was used to calculate the stiffness of the sample [82]. The characteristic force-displacement curves were measured to determine the Young's modulus of the membrane. Mikczinski et al. [83] fixed a 2D nanopaper made of microfibrillated cellulose at both ends and drove it against a capacitive force sensor to perform nanoindentation experiments to determine its Young's modulus. In Fig. 11.18d, Fatikow et al used a self-sensing piezoresistive AFM cantilever to conduct nanoindentation experiments on graphene flakes which were transferred and soldered to be suspended on a standard aluminum TEM grid for mechanical characterization [84, 85].

The schematic in Fig. 11.18e illustrates tensile experiments for piezoresistivity characterization of a suspended Si nanowire performed by Lugstein et al. The two ends of the Si nanowire under test were anchored with a suspended cantilever and an insulated Si pad respectively. The probe in green is used to apply stress to the freestanding cantilever, generating strain on the nanowire. The two probes in red were applied to form electrical connections to measure the resistance changes [86]. Zimmermann et al. [87] used an L-shaped four-point probe fabricated using FIB for electrical characterization of a graphene flake fabricated by mechanical exfoliation, illustrated in Fig. 11.18f).

11.5.2 Construction of Nanoscale Devices

Not only has nanorobotic manipulation become an enabling technology for characterizing the properties of nanomaterials, but it has also been further applied for the construction of nanoscale building blocks and devices.

Figure 11.19a shows individual gold nanowires that were picked and placed by two probes, and then welded and assembled to form a nanoscale pattern "NANO" by Peng et al. [63]. A 3D letter was N formed on a MWCNT by oxygen assisted cutting, bending, and manipulation for assembly, as shown in Fig. 11.19b [88]. Oxygen gas can be introduced into the vicinity of a CNT in SEM through a glass nozzle while irradiating it with a low-energy electron beam to cut and bend the CNT in order to from 3D structures, as shown in [89, 90]. The cutting mechanism is opposite to the deposition method. By exposing a CNT to an electron beam, carbon heats up

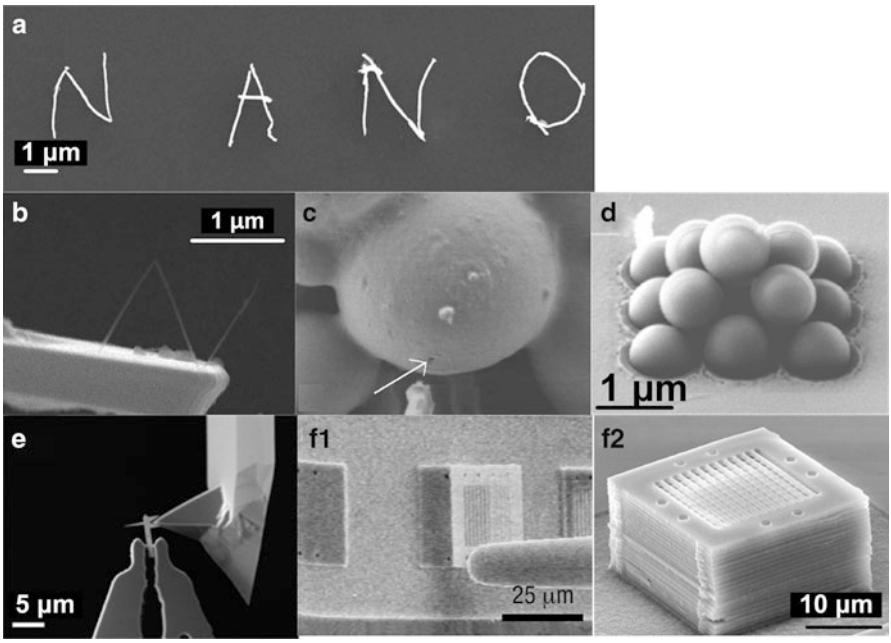


Fig. 11.19 Assembly and construction for nanoscale blocks and devices. (**a**) Nanowires were picked and placed to spell "NANO" [63]. (**b**) A MWCNT was used to form a 3D letter "N" using oxygen assisted fabrication [88]. (**c**) Holes are pierced in polymer microcapsules to release medical agents [92]. (**d**) Microspheres are stacked using two cooperative manipulators [25]. (**e**) A tool is mounted on an AFM cantilever [64]. (**f1**–**f2**) Photonic plates are removed from a substrate, picked and placed on a structure to produce a photonic crystal [93]

and ionizes, reacting with oxygen to generate carbon monoxide or carbon dioxide to cut the CNT. Yuzvinsky et al. have also demonstrated CNT cutting with water vapor [91] and have also achieved partial cuts in CNT, making hinge-like structures. By increasing the acceleration voltage of the electron beam, placing the gas nozzle further from the sample, or decreasing irradiation time, CNTs can be bent instead of cut [88]. Nanoscale holes were produced by piercing of polymer microcapsules by precise control of a microtool of microelectrode made from glass capillary [92], as indicated by the arrow in Fig. 11.19c, to release encapsulated medical active agents for targeted drug delivery. Tiemerding et al. created a 3D pyramidal structure consisting of stacked silica spheres via cooperative manipulation of two probes, as shown in Fig. 11.19d. An electrochemically etched tungsten probe with a high-aspect ratio tip and a modified piezoresistive AFM cantilever with a spherical adhesion pocket tip were fabricated to pick and place the spheres [25]. In Fig. 11.19e, a nanoscale structure was manipulated by a microgripper to approach an AFM cantilever, and then soldered and assembled onto an AFM tip using EBID to construct new tools [64]. For a photonics application, Aoki et al. separated a photonic plate from a substrate after breaking the connection bridge with a probe,

then picked and transferred the plate onto a substructure, as shown in Fig. 11.19f1. More plates were stacked and aligned using the same procedure to form a 3D photonic crystal, shown in Fig. 11.19f2, showing potential for the production of optical wavelength photonic crystal devices [93].

11.5.3 Cell Nanosurgery in ESEM

Environmental scanning electron microscopes (ESEMs) enable high resolution (e.g., 3.5 nm) and real-time imaging and observation of water-containing samples and electrically insulating materials using a specially built secondary electron detector [94]. Both TEMs and conventional SEMs work under high-vacuum conditions; therefore, they cannot directly observe water-containing samples (e.g., biological cells in liquids). ESEM overcomes these limitations to allow manipulation and characterization of biological, water-containing samples [9, 95].

A nanomanipulator system was constructed inside an ESEM with three units, two units for manipulation with 7 DOFs and one cooling stage to hold samples and control their temperature. This system can simultaneously perform real-time observation and manipulation of biological samples. Commercial AFM cantilevers were modified by FIB to fabricate various tools for manipulation and cell characterization. A soft buckling nano-needle was made based on the modification of a commercial AFM cantilever to perform indentation experiments, as shown in Fig. 11.20a. Cell stiffness can be calculated after imaging processing [94, 95]. A nano-fork and nano/micro-putter were fabricated by FIB to perform cell lifting and side pushing experiments for investigating adhesion between the cell and substrate [96, 97], as shown in Fig. 11.20b, c. A nano-picker was also constructed to perform cell–cell adhesion force measurements [98]. An in-situ single cell cutting experiment was performed with a nano-knife, and the cutting force and sample slice angle were evaluated by the buffering beam deformation [54], shown in Fig. 11.20d. The stiffness of fabricated tools can be calibrated experimentally using nanomanipulation techniques, by comparing the deflection of the tools to a reference cantilever with a known stiffness [96]. A dual nano-probe was also fabricated from a standard piezoresistive cantilever to conduct cell electrical measurements to determine viability [99], shown in Fig. 11.20e. An AFM cantilever was modified by FIB into a sharp tip with a neck to deliver fluorescent materials into Caenorhabditis elegans, and can be broken at the neck point as an embedded nanoscale injector after insertion into the biological sample [100], as shown in Fig. 11.20f).

11.5.4 Hybrid Nanorobotic Manipulation in TEM and SEM

TEMs provide real-time, sub-nanometer imaging. A hybrid nanorobotic manipulation system integrated with an 8-DOF SEM manipulator subsystem and a 6-DOF

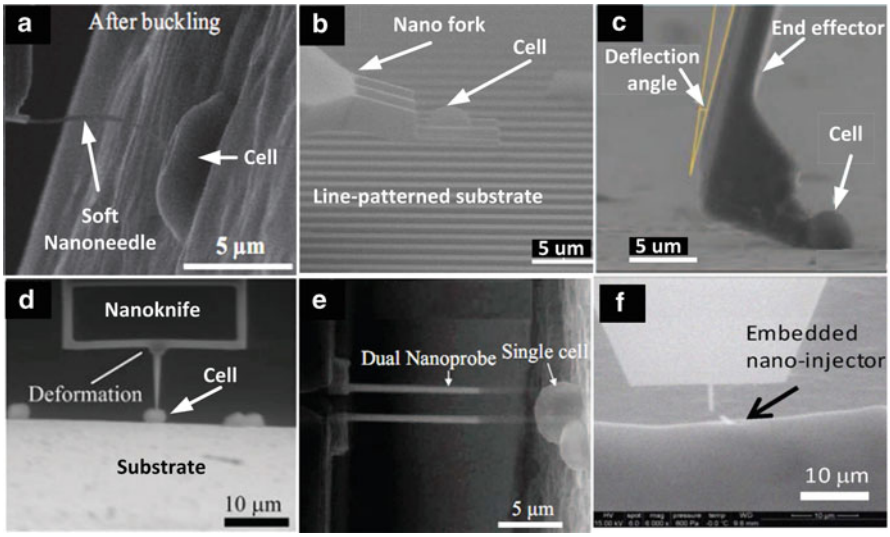


Fig. 11.20 Various tools for single cell measurement under ESEM. (**a**) Cell stiffness testing. Adapted from [94]. (**b**–**c**) Nano fork and nano putter for testing cell adhesion. Adapted from [96, 97]. (**d**) Nano knife for cutting single cells. Adapted from [54]. (**e**) Testing cell viability using a dual nanoprobe. Adapted from [99]. (**f**) Nanoinjection of C. Elegans. Adapted from [100]

TEM manipulator inside a SEM and TEM, respectively, was constructed to take advantages of these two types of electron microscopes. This hybrid system is capable of performing complicated sample preparation by taking advantage of SEM's large workspace and high imaging resolution of TEM [5, 6]. A schematic of the SEM and TEM hybrid system is shown in Fig. 11.21. The TEM manipulator is prepared on the SEM manipulator. An example application for CNT mechanical characterization was performed as a demonstration. A CNT was pre-manipulated for positioning the sample onto the TEM manipulator inside the SEM, followed by stimulating the CNT for measurement with the TEM manipulator inside the TEM. The experiment showed improved accuracy in measuring the Young's modulus of the CNT [6].

11.5.5 Automated Tasks

Manual nanomanipulation by joystick is time-consuming and skill dependent. Success rate can vary significantly due to poor reproducibility and inconsistency among operators. There is a steep learning curve, and intense focus is required for operators to avoid damaging samples or the fragile end effectors. Automated manipulation and assembly tasks using visual tracking and servoing techniques can be more reliable, efficient, and reproducible than manual operation.

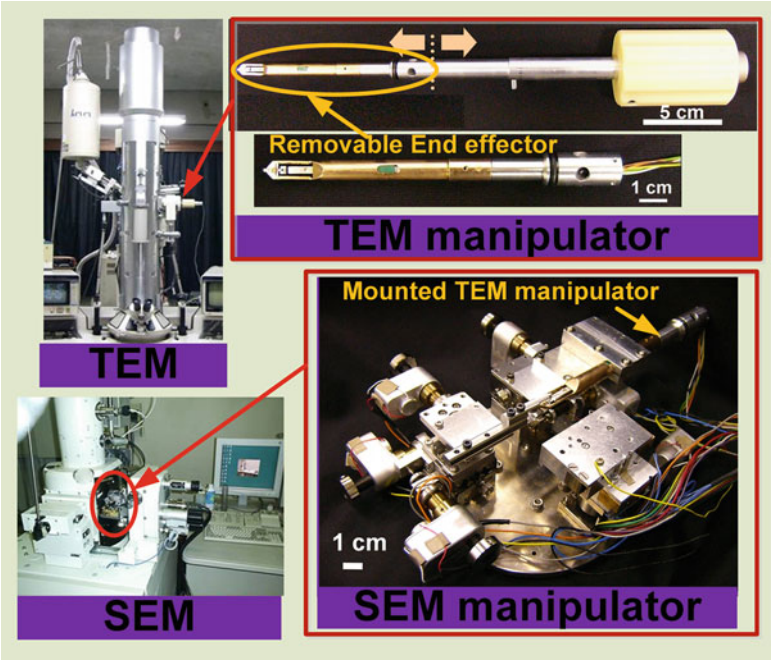


Fig. 11.21 Nanorobotic manipulators inside TEM and SEM. Adapted from [5].

11.5.5.1 Pick/Place and Assembly

Ye et al. [10] used a nanomanipulation system to automatically pick and place an individual silicon nanowire and then transfer it to a MEMS device for characterization, shown in Fig. 11.22.

The system selects a nanowire from the growing substrate using image processing, as shown in Fig. 11.22a. To pick up the nanowire, the probe moves to contact the nanowire, and then EBID is used to solder the nanowire onto the probe at the contact point, as shown in Fig. 11.22b. To place the nanowire, the probe moves the nanowire to the desired location. The nanowire is EBID soldered in place as shown in Fig. 11.22c, d. Both mechanical and electrical characterization experiments were performed on the MEMS device following this pick-place procedure [10, 101]. Zhang et al. [102] demonstrated a method for nanodevice construction by controlling a nanomanipulation system to perform nanowire selection.

11.5.5.2 Electrical Probing

Nanoprobing inside the SEM allows nanoprobe tips to be precisely positioned on nanostructures for electrical characterization, which is important for failure analysis,

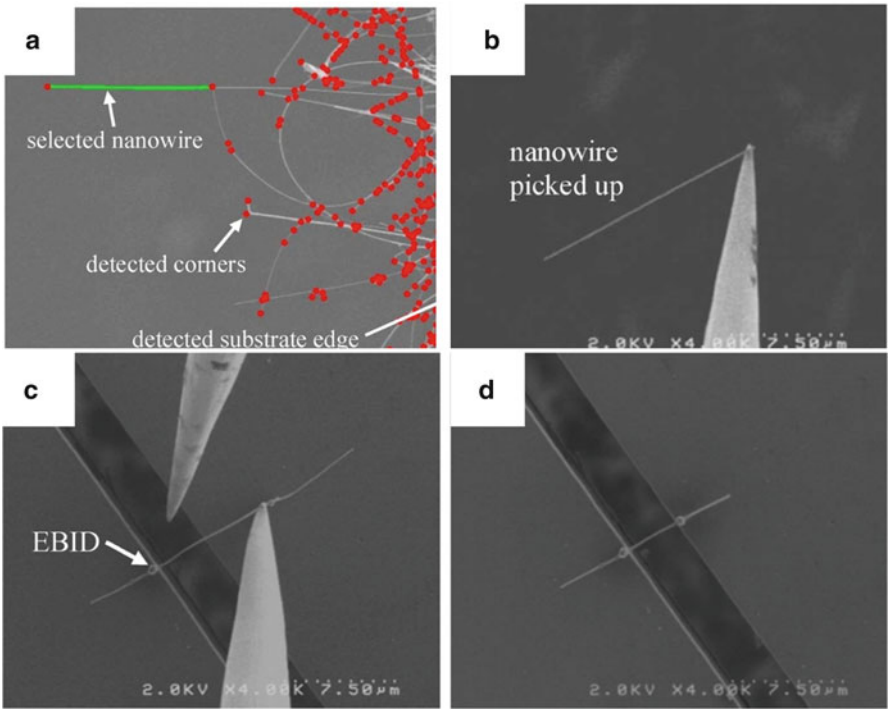


Fig. 11.22 Pick-place of nanowire. (**a**) Image processing software automatically detects and selects a nanowire. (**b**) Probe removes nanowire from substrate. (**c**) Nanowire is EBID soldered to one side of the gap. (**d**) Nanowire is EBID soldered to other side of the gap. Electrical breakdown between probe 1 and nanowire releases the nanowire from the probe. Adapted from [10]

quality control, and process development in the semiconductor industry. Automated nanoprobing was performed as shown in Fig. 11.23 to probe nanostructures on a SEM metrology chip. Algorithms were developed to de-noise the SEM image and compensate for drift in real time, enabling robust visual tracking and servo control for probing nanostructures. Using the SEM as a vision sensor for automation and contact detection [24], Ru et al. landed four probes on a silicon nanowire to perform four-point probing of single nanowires with a 100 % success rate over 50 trials [58, 103]. A feedforward controller was applied to improve the response time for controlling multiple nanomanipulators [58, 103]. Besides nanowires, this process can also be used to probe electrical properties of nanostructures [59] or other nanomaterials. To enable faster and more accurate positioning without relying on SEM visual feedback, Zhou et al. developed a strain gauge-based position feedback system, which demonstrated a closed-loop positioning accuracy better than 3 nm [8].

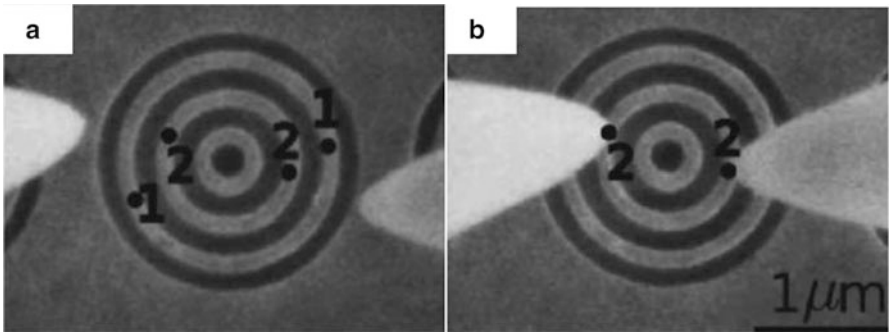


Fig. 11.23 Automated probing of nanostructure and silicon nanowire. (**a**) User identifies points to probe from SEM image. (**b**) Nanomanipulator probes the specified points. Adapted from [58]

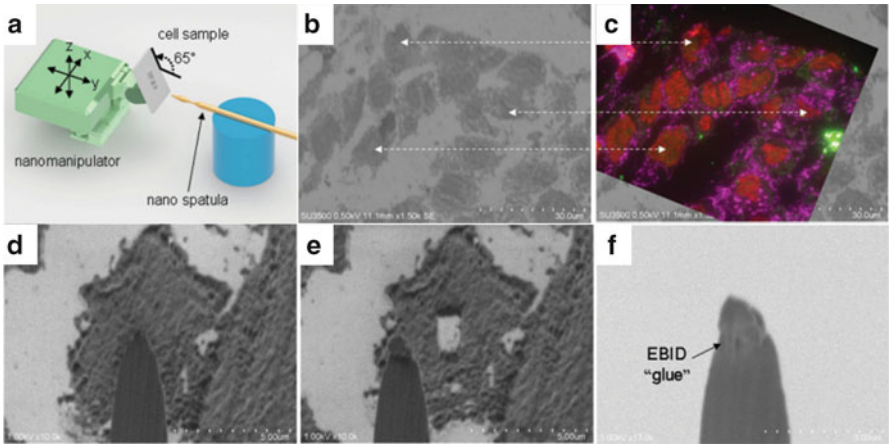


Fig. 11.24 Chromatin extraction process. (**a**) DNA extraction setup in SEM. (**b**) and (**c**) SEM and fluorescence imaging correlation. (**d**) Land probe on selected cell. (**e**) Press nano-spatula into surface and scrape off cell fragment. (**f**) Lift nano-spatula from sample and EBID "glue" cell fragment to nano-spatula [104]

11.5.5.3 DNA Extraction

Nanomanipulation was demonstrated to perform cell nanosurgery inside SEM. Chen et al. [104] utilized cell nanosurgery to infer functions of cell sub-nuclear structures by examining nearby gene loci by using a "nano-spatula" tool to scoop up a single chromatin as shown in Fig. 11.24. To facilitate the process, SEM images were correlated with fluorescence microscope images to determine target locations to scoop [105, 106]. Because gene locations within the nucleus are not random, this technique can be used to discover functions and gene regulation mechanisms of sub-nuclear structures.

11.6 Conclusion

Nanoscale laboratories inside SEM have been constructed and integrated with AFMs, FIBs, optical microscopes, multi-tool changers, and other new nanoscale fabrication techniques. These technologies provide a foundation for bottom-up fabrication and a powerful platform for characterizing nanoscale materials, forming nanoscale building blocks, and assembling nanodevice prototypes. Technologies such as EBID and FIB permit micro/nanodevice prototypes to be directly machined or deposited, shortening the development time. These tools are expected to help make new discoveries in biology, material science, and semiconductor industries. Technically, new automation techniques will continue to be developed and will emerge to achieve enhanced efficiency and reproducibility.

References

1. D.M. Eigler, E.K. Schweizer, Positioning single atoms with a scanning tunnelling microscope. Nature **344**(6266), 524–526 (1990)
2. S. Kim, F. Shafiei, D. Ratchford, X. Li, Controlled AFM manipulation of small nanoparticles and assembly of hybrid nanostructures. Nanotechnology **22** (11), 115301 (2011)
3. U. Mick, M. Weigel-Jech, S. Fatikow, Robotic workstation for AFM-based nanomanipulation inside an SEM, in *2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)* (2010), pp. 854–859
4. U. Mick, V. Eichhorn, T. Wortmann, C. Diederichs, S. Fatikow, Combined nanorobotic AFM/SEM system as novel toolbox for automated hybrid analysis and manipulation of nanoscale objects, in *2010 IEEE International Conference on Robotics and Automation (ICRA)* (2010), pp. 4088–4093
5. T. Fukuda, M. Nakajima, P. Liu, M.R. Ahmad, Bringing the nanolaboratory inside electron microscopes. Nanotechnol. Mag. IEEE **2**(2), 18–31 (2008)
6. M. Nakajima, F. Arai, T. Fukuda, In situ measurement of Young's modulus of carbon nanotubes inside a TEM through a hybrid nanorobotic manipulation system. IEEE Trans. Nanotechnol. **5**(3), 243–248 (2006)
7. T. Fukuda, M. Nakajima, M. Ahmad, Y. Shen, M. Kojima, Micro- and nanomechatronics. IEEE Ind. Electron. Mag. **4**(4), 13–22 (2010)
8. C. Zhou, Z. Gong, B.K. Chen, M. Tan, Y. Sun, Closed-loop controlled nanoprobing inside SEM, in *2014 IEEE 14th International Conference on Nanotechnology (IEEE-NANO)* (2014), pp. 45–48
9. T. Fukuda, M. Nakajima, H. Tajima, Y. Shen, T. Yue, Micro-nanomanipulation system toward biological cell analysis and assembly, in *2012 First International Conference on Innovative Engineering Systems (ICIES)* (2012), pp. 31–36
10. X. Ye, Y. Zhang, C. Ru, J. Luo, S. Xie, Y. Sun, Automated pick-place of silicon nanowires. IEEE Trans. Autom. Sci. Eng. **10**(3), 554–561 (2013)
11. S. Fahlbusch, S. Mazerolle, J.-M. Breguet, A. Steinecker, J. Agnus, R. Pérez, J. Michler, Nanomanipulation in a scanning electron microscope. J. Mater. Process. Technol. **167**(2–3), 371–382 (2005)
12. S. Fatikow, V. Eichhorn, M. Bartenwerfer, F. Krohs, Nanorobotic AFM/SEM/FIB system for processing, manipulation and characterization of nanomaterials, in *2013 IEEE 18th Conference on Emerging Technologies Factory Automation (ETFA)* (2013), pp. 1–4

13. J.I. Goldstein, D.E. Newbury, P. Echlin, D.C. Joy, A.D.R. Jr, C.E. Lyman, C. Fiori, E. Lifshin, Introduction, in *Scanning Electron Microscopy and X-Ray Microanalysis*, Springer US, 1992, pp. 1–19
14. "University of Glasgow:: Schools:: School of Geographical and Earth Sciences," *University of Glasgow:: Glasgow, Scotland, UK*. [Online]. Available: http://www.gla.ac.uk/schools/ges/research/researchfacilities/isaac/services/scanningelectronmicroscopy/imaging/secondaryelectronseandbackscatteredelectronbse/
15. J. Portilla, V. Strela, M.J. Wainwright, E.P. Simoncelli, Image denoising using scale mixtures of Gaussians in the wavelet domain. IEEE Trans. Image Process. **12**, 1338–1351 (2003)
16. S.M. Smith, J.M. Brady, SUSAN—a new approach to low level image processing. Int. J. Comput. Vis. **23**(1), 45–78 (1997)
17. F. Catté, P.-L. Lions, J.-M. Morel, T. Coll, Image selective smoothing and edge detection by nonlinear diffusion. SIAM J. Numer. Anal. **29**(1), 182–193 (1992)
18. C. Vogel, M. Oman, Iterative methods for total variation denoising. SIAM J. Sci. Comput. **17**(1), 227–238 (1996)
19. N. Ouarti, B. Sauvet, S. Régnier, High quality real-time video with scanning electron microscope using total variation algorithm on a graphics processing unit. Int. J. Optomechatronics **6**(2), 163–178 (2012)
20. A. Buades, B. Coll, A non-local algorithm for image denoising, in *In CVPR* (2005), pp. 60–65
21. F. Goudail, P. Réfrégier, *Statistical Image Processing Techniques for Noisy Images* (Springer US, Boston, 2004)
22. Z. Gong, B.K. Chen, J. Liu, Y. Sun, Automated nanoprobing under scanning electron microscopy, in *2013 IEEE International Conference on Robotics and Automation (ICRA)* (2013), pp. 1433–1438
23. T. Tiemerding, C. Diederichs, S. Zimmermann, S. Fatikow, Closing the loop: high-speed visual servoing and control of a commercial nanostage inside the SEM, in *2013 International Conference on Manipulation, Manufacturing and Measurement on the Nanoscale (3M-NANO)* (2013), pp. 35–39
24. C. Ru, S. To, Contact detection for nanomanipulation in a scanning electron microscope. Ultramicroscopy **118**, 61–66 (2012)
25. T. Tiemerding, S. Zimmermann, S. Fatikow, Robotic dual probe setup for reliable pick and place processing on the nanoscale using haptic devices, in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014)* (2014), pp. 892–897
26. V. Eichhorn, S. Fatikow, T. Wich, C. Dahmen, T. Sievers, K.N. Andersen, K. Carlson, P. Bøggild, Depth-detection methods for microgripper based CNT manipulation in a scanning electron microscope. J. Micro-Nano Mechatron. **4**(1–2), 27–36 (2008)
27. T. Kasaya, H.T. Miyazaki, S. Saito, K. Koyano, T. Yamaura, T. Sato, Image-based autonomous micromanipulation system for arrangement of spheres in a scanning electron microscope. Rev. Sci. Instrum. **75**(6), 2033–2042 (2004)
28. A. Cvetanovic, A. Cvetanovic, A. Deutschinger, I. Giouroudi, W. Brenner, Design of a novel visual and control system for the prevention of the collision during the micro handling in a SEM chamber. Microelectron. Eng. **87**(2), 139–143 (2010)
29. E. Ribeiro, M. Shah, Computer vision for nanoscale imaging. Mach. Vis. Appl. **17**(3), 147–162 (2006)
30. S. Fatikow, T. Wich, T. Sievers, M. Jähnisch, V. Eichhorn, J. Mircea, H. Hülsen, C. Stolle, Automatic nanohandling station inside a scanning electron microscope. Proc. Inst. Mech. Eng. B J. Eng. Manuf. **222**(1), 117–128 (2008)
31. M. Jähnisch, S. Fatikow, 3D vision feedback for nanohandling monitoring in a scanning electron microscope. Int. J. Optomechatronics **1**, 4–26 (2007)
32. V. Eichhorn, S. Fatikow, T. Wortmann, C. Stolle, C. Edeler, D. Jasper, O. Sardan, P. Boggild, G. Boetsch, C. Canales, R. Clavel, NanoLab: a nanorobotic system for automated pick-and-place handling and characterization of CNTs, in *IEEE International Conference on Robotics and Automation, 2009. ICRA '09* (2009), pp. 1826–1831

33. V. Eichhorn, S. Fatikow, Ch. Dahmen, Ch. Edeler, C. Stolle, D. Jasper (Presenter), Automated microfactory inside a scanning electron microscope, in *Proc. of the 6th International Workshop on Microfactories (IWMF)*, Evanston, 5–7 October 2008
34. W. Denk, H. Horstmann, Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. PLoS Biol.**2**(11), e329 (2004)
35. B. Andres, U. Koethe, T. Kroeger, M. Helmstaedter, K.L. Briggman, W. Denk, F.A. Hamprecht, 3D segmentation of SBFSEM images of neuropil by a graphical model over supervoxel boundaries. Med. Image Anal. **16**(4), 796–805 (2012)
36. V. Marx, Neurobiology: brain mapping in high resolution. Nature **503**(7474), 147–152 (2013)
37. "Nanotechnology at Zyvex," *Nanotechnology at Zyvex*, 2010. [Online]. Available: http://www.zyvex.com/Products/S100_Features.html
38. J. Zuverza, "NanoBot® Nanomanipulator System from Xidex," *Xidex Nanobot® Nanomanipulator and Nanotechnology Products*, 2011. [Online]. Available: http://xidex.com/products/nanobot.html
39. "Nanomanipulators - OmniProbe 400 - Oxford Instruments," *Home Page - Oxford Instruments*, 2015. [Online]. Available: http://www.oxford-instruments.com/products/nanomanipulation-nanofabrication/nanomanipulator/omniprobe-400
40. B. K. Chen, LifeForce Nanomanipulation System: LF-2000
41. "Micromanipulator miBot™ BT-11 | Imina Technologies SA," *Imina Technologies SA | High Precision Robots for Microscopes*, 2014. [Online]. Available: http://imina.ch/technology
42. "Kleindiek Nanotechnik: Electron Microscopy," *Kleindiek Nanotechnik*, 2015. [Online]. Available: http://www.nanotechnik.com/mm3a-em.html
43. P. Russel, D. Batchelor, SEM and AFM: complementary techniques for surface investigations. Microsc. Anal. **2001**, 9–14 (2001)
44. "NSOM/SNOM, TERS, Low Temperature AFM & SPM Solutions from Nanonics." [Online]. Available: http://www.nanonics.co.il/
45. "Klocke Nanotechnik, Aachen." [Online]. Available: http://nanomotor.de/index.html
46. "Hybrid AFM SEM Combination, the DME BRR Microscope," *Manufacturer of Atomic Force Microscopes (AFM), Scanning Tunneling Microscopes (STM)*, 2013. [Online]. Available: http://www.dme-spm.com/remafm.html
47. "Test of camera modules, autocollimator, MTF system - TRIOPTICS," 2015. [Online]. Available: http://www.trioptics.com/
48. "Atomic force microscope / electronic - attoAFM/SEM - attocube systems AG," *DirectIndustry - The Virtual Industrial Exhibition: sensor - automation - motor - pump - handling - packaging*, 2015. [Online]. Available: http://www.directindustry.com/prod/attocube-systems-ag/atomic-force-microscope-electronic-50096-1321683.html
49. F. Long, C. Wang, M. Lü, F. Zhang, J. Sun, J. Hu, Optimizing single DNA molecules manipulation by AFM. J. Microsc. **243**(2), 118–123 (2011)
50. T. Hertel, R. Martel, P. Avouris, Manipulation of individual carbon nanotubes and their interaction with surfaces. J. Phys. Chem. B **102**(6), 910–915 (1998)
51. G. Li, N. Xi, H. Chen, C. Pomeroy, M. Prokos, 'Videolized' atomic force microscopy for interactive nanomanipulation and nanoassembly. IEEE Trans. Nanotechnol. **4**(5), 605–615 (2005)
52. M.R. Ahmad, M. Nakajima, S. Kojima, M. Homma, T. Fukuda, In-situ single cell mechanical characterization of W303 Yeast cells inside Environmental-SEM, in *7th IEEE Conference on Nanotechnology, 2007. IEEE-NANO 2007* (2007), pp. 1022–1027
53. Y. Shen, M. Nakajima, M. Homma, T. Fukuda, Auto nanomanipulation system for single cell mechanical property characterization inside an environmental SEM, in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2012), pp. 646–651
54. Y. Shen, M. Nakajima, Z. Yang, S. Kojima, M. Homma, T. Fukuda, Design and characterization of nanoknife with buffering beam for in situ single-cell cutting. Nanotechnology **22**(30), 305701 (2011)

55. V. Thanh, S.A. Chizhik, T. Xuan, N. Trong, V.V. Chikunov, Tuning fork scanning probe microscopes – applications for the nano-analysis of the material surface and local physico-mechanical properties, in *Scanning Probe Microscopy-Physical Property Characterization at Nanoscale*, ed. by V. Nalladega (InTech, 2012)
56. F.J. Giessibl, Principles and Applications of the qPlus Sensor, in *Noncontact Atomic Force Microscopy*, ed. by S. Morita, F.J. Giessibl, R. Wiesendanger (Springer, Berlin, Heidelberg, 2009), pp. 121–142
57. T. Akiyama, U. Staufer, N. de Rooij, Self-sensing and self-actuating probe based on quartz tuning fork combined with microfabricated cantilever for dynamic mode atomic force microscopy. Appl. Surf. Sci. **210**(1–2), 18–21 (2003)
58. C. Ru, Y. Zhang, Y. Sun, Y. Zhong, X. Sun, D. Hoyle, I. Cotton, Automated four-point probe measurement of nanowires inside a scanning electron microscope. IEEE Trans. Nanotechnol. **10**(4), 674–681 (2011)
59. Z. Gong, B.K. Chen, J. Liu, Y. Sun, Robotic probing of nanostructures inside scanning electron microscopy. IEEE Trans. Robot. **30**(3), 758–765 (2014)
60. L. Dong, B.J. Nelson, Tutorial – robotics in the small part II: nanorobotics. IEEE Robot. Autom. Mag. **14**(3), 111–121 (2007)
61. T. Fukuda, F. Arai, L. Dong, Assembly of nanodevices with carbon nanotubes through nanorobotic manipulations. Proc. IEEE **91**(11), 1803–1818 (2003)
62. B.K. Chen, Y. Zhang, D.D. Perovic, Y. Sun, MEMS microgrippers with thin gripping tips. J. Micromech. Microeng. **21**(10), 105004 (2011)
63. Y. Peng, T. Cullis, B. Inkson, Bottom-up nanoconstruction by the welding of individual metallic nanoobjects using nanoscale solder. Nano Lett. **9**(1), 91–96 (2009)
64. R.T.R. Kumar, S.U. Hassan, O.S. Sukas, V. Eichhorn, F. Krohs, S. Fatikow, P. Boggild, Nanobits: customizable scanning probe tips. Nanotechnology **20**(39), 395703 (2009)
65. M. Nakajima, T. Kawamoto, M. Kojima, T. Fukuda, Nanotool exchanger system using low-melting metal under environmental SEM, in *2011 International Symposium on Micro-NanoMechatronics and Human Science (MHS)*, 2011, pp. 69–74
66. A. Garetto, C. Baur, J. Oster, M. Waiblinger, K. Edinger, Advanced process capabilities for electron beam based photomask repair in a production environment, in *Proc. SPIE 7122* (2008), p. 7122k
67. L. van Kouwen, A. Botman, C.W. Hagen, Focused electron-beam-induced deposition of 3 nm dots in a scanning electron microscope. Nano Lett. **9**(5), 2149–2152 (2009)
68. M. Huth, F. Porrati, C. Schwalb, M. Winhold, R. Sachser, M. Dukic, J. Adams, G. Fantner, Focused electron beam induced deposition: a perspective. Beilstein J. Nanotechnol. **3**, 597–619 (2012)
69. I. Utke, S. Moshkalev, P. Russell, *Nanofabrication Using Focused Ion and Electron Beams: Principles and Applications* (Oxford University Press, New York, 2012)
70. V. Scheuer, H. Koops, T. Tschudi, Electron beam decomposition of carbonyls on silicon. Microelectron. Eng. **5**(1–4), 423–430 (1986)
71. S.J. Randolph, J.D. Fowlkes, P.D. Rack, Focused, nanoscale electron-beam-induced deposition and etching. Crit. Rev. Solid State Mater. Sci. **31**(3), 55–89 (2006)
72. "Nanobot Nanomanipulator - Parallel Gas Injection System," *Xidex Nanobot Nanomanipulator and Nanotechnology Products*, 2011. [Online]. Available: http://xidex.com/products/parallel-gas-injection-system.html
73. "Gas Injection Systems: OmniGIS II – Oxford Instruments," 2015. [Online]. Available: http://www.oxford-instruments.com/products/nanomanipulation-nanofabrication/gas-injection-system/omnigis-ii
74. "GIS-5: Multi-lines Gas Injection System," *Orsay Physics*, 2015. [Online]. Available: http://www.orsayphysics.com/product-gis-5.html
75. Z. Yang, M. Nakajima, Y. Saito, Y. Ode, T. Fukuda, Isolated high-purity platinum nanowire growth via field emission from a multi-walled carbon nanotube. Appl. Phys. Express **4**(3), 035001 (2011)

76. Z. Yang, M. Nakajima, Y. Ode, T. Fukuda, Tungsten/platinum hybrid nanowire growth via field emission using nanorobotic manipulation. J. Nanotechnol. **2011**, e386582 (2011)
77. Z. Yang, M. Nakajima, Y. Ode, T. Fukuda, Nanoassembly of pH sensor nanoprobe by multiple-metallic nanowires, in *2010 10th IEEE Conference on Nanotechnology (IEEE-NANO)* (2010), pp. 352–355
78. Z. Yang, M. Nakajima, Y. Ode, Z. Zhang, T. Fukuda, Fabrication and evaluation of nano probe pH sensor based on nanorobotic manipulation, in *2010 International Symposium on Micro-NanoMechatronics and Human Science (MHS)*, 2010, pp. 284–289
79. Y. Zhu, Q. Qin, F. Xu, F. Fan, Y. Ding, T. Zhang, B.J. Wiley, Z.L. Wang, Size effects on elasticity, yielding, and fracture of silver nanowires: in situ experiments. Phys. Rev. B **85**(4), 045443 (2012)
80. D.J. Bell, L. Dong, B.J. Nelson, M. Golling, L. Zhang, D. Grützmacher, Fabrication and characterization of three-dimensional InGaAs/GaAs nanosprings. Nano Lett. **6**(4), 725–729 (2006)
81. J.-O. Abrahamians, B. Sauvet, J. Polesel-Maris, R. Braive, S. Regnier, A nanorobotic system for in situ stiffness measurements on membranes. IEEE Trans. Robot. **30**(1), 119–124 (2014)
82. A. Castellanos-Gomez, N. Agraït, G. Rubio-Bollinger, Dynamics of quartz tuning fork force sensors used in scanning probe microscopy. Nanotechnology **20**(21), 215502 (2009)
83. M.R. Mikczinski, G. Josefsson, G. Chinga-Carrasco, E.K. Gamstedt, S. Fatikow, Nanorobotic testing to assess the stiffness properties of nanopaper. IEEE Trans. Robot. **30**(1), 115–119 (2014)
84. S. Zimmermann, V. Eichhorn, S. Fatikow, Nanorobotic transfer and characterization of graphene flakes, in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2012), pp. 640–645
85. S. Zimmermann, T. Tiemerding, T. Li, W. Wang, Y. Wang, S. Fatikow, Automated mechanical characterization of 2-D materials using SEM based visual servoing. Int. J. Optomechatronics **7**(4), 283–295 (2013)
86. A. Lugstein, M. Steinmair, A. Steiger, H. Kosina, E. Bertagnolli, Anomalous piezoresistance effect in ultrastrained silicon nanowires. Nano Lett. **10**(8), 3204–3208 (2010)
87. S. Zimmermann, S.A.G. Barragan, S. Fatikow, Nanorobotic processing of graphene: a platform tailored for rapid prototyping of graphene-based devices. IEEE Nanotechnol. Mag. **8**(3), 14–19 (2014)
88. P. Liu, K. Kantola, T. Fukuda, F. Arai, Nanoassembly of nanostructures by cutting, bending and soldering of carbon nanotubes with electron beam. J. Nanosci. Nanotechnol. **9**(5), 3040–3045 (2009)
89. P. Liu, F. Arai, T. Fukuda, Cutting of carbon nanotubes assisted with oxygen gas inside a scanning electron microscope. Appl. Phys. Lett. **89**(11), 113104 (2006)
90. P. Liu, F. Arai, T. Fukuda, Nanofabrication of carbon nanotubes assisted with oxygen gas, in *Sixth IEEE Conference on Nanotechnology, 2006. IEEE-NANO 2006*, vol. 2 (2006), pp. 540–543
91. T.D. Yuzvinsky, A.M. Fennimore, W. Mickelson, C. Esquivias, A. Zettl, Precision cutting of nanotubes with a low-energy electron beam. Appl. Phys. Lett., **86**(5), 053109 (2005)
92. E. Meyer, H.-G. Braun, Micro- and nanomanipulation inside the SEM. J. Phys. Conf. Ser. **126**, 012074 (2008)
93. K. Aoki, H.T. Miyazaki, H. Hirayama, K. Inoshita, T. Baba, K. Sakoda, N. Shinya, Y. Aoyagi, Microassembly of semiconductor three-dimensional photonic crystals. Nat. Mater. **2**(2), 117–121 (2003)
94. M.R. Ahmad, M. Nakajima, S. Kojima, M. Homma, T. Fukuda, The effects of cell sizes, environmental conditions, and growth phases on the strength of individual W303 yeast cells inside ESEM. IEEE Trans. Nanobioscience **7**(3), 185–193 (2008)
95. M.R. Ahmad, M. Nakajima, S. Kojima, M. Homma, T. Fukuda, In situ single cell mechanics characterization of yeast cells using nanoneedles inside environmental SEM. IEEE Trans. Nanotechnol. **7**(5), 607–616 (2008)

96. M.R. Ahmad, M. Nakajima, M. Kojima, S. Kojima, M. Homma, T. Fukuda, Nanofork for single cells adhesion measurement via ESEM-nanomanipulator system. IEEE Trans. Nanobioscience **11**(1), 70–78 (2012)
97. Y. Shen, M. Nakajima, S. Kojima, M. Homma, M. Kojima, T. Fukuda, Single cell adhesion force measurement for cell viability identification using an AFM cantilever-based micro putter. Meas. Sci. Technol. **22**(11), 115802 (2011)
98. Y. Shen, M. Nakajima, S. Kojima, M. Homma, T. Fukuda, Study of the time effect on the strength of cell-cell adhesion force by a novel nano-picker. Biochem. Biophys. Res. Commun. **409**(2), 160–165 (2011)
99. M.R. Ahmad, M. Nakajima, M. Kojima, S. Kojima, M. Homma, T. Fukuda, Instantaneous and quantitative single cells viability determination using dual nanoprobe inside ESEM. IEEE Trans. Nanotechnol. **11**(2), 298–306 (2012)
100. T. Hirano, M. Nakajima, M. Kojima, N. Hisamoto, M. Homma, T. Fukuda, Selective nano-injection using nano-probe based on nanomanipulation under hybrid microscope, in *2011 International Symposium on Micro-NanoMechatronics and Human Science (MHS)* (2011), pp. 216–221
101. Y. Zhang, X. Liu, C. Ru, Y.L. Zhang, L. Dong, Y. Sun, Piezoresistivity characterization of synthetic silicon nanowires using a MEMS device. J. Microelectromech. Syst. **20**(4), 959–967 (2011)
102. Y.L. Zhang, J. Li, S. To, Y. Zhang, X. Ye, L. You, Y. Sun, Automated nanomanipulation for nanodevice construction. Nanotechnology **23**(6), 065304 (2012)
103. C. Ru, Y. Zhang, Y. Sun, Y. Zhong, X. Sun, D. Hoyle, I. Cotton, Automated four-point probe measurement of nanowires inside a scanning electron microscope, in *2010 IEEE Conference on Automation Science and Engineering (CASE)* (2010), pp. 533–538
104. B.K. Chen, D. Anchel, Z. Gong, R. Cotton, R. Li, Y. Sun, D.P. Bazett-Jones, Gene organization: nano-dissection and sequencing of DNA at single sub-nuclear structures (Small 16/2014). Small **10**(16), 3266–3266 (2014)
105. Z. Gong, B.K. Chen, J. Liu, C. Zhou, D. Anchel, X. Li, D.P. Bazett-Jones, Y. Sun, Correlative microscopy for nanomanipulation of sub-cellular structures, in *2014 IEEE International Conference on Robotics and Automation (ICRA)* (2014), pp. 5209–5214
106. Z. Gong, B.K. Chen, J. Liu, C. Zhou, D. Anchel, X. Li, J. Ge, D.P. Bazett-Jones, Y. Sun, Fluorescence and SEM correlative microscopy for nanomanipulation of subcellular structures. Light Sci. Appl. **3**(11), e224 (2014)

Chapter 12
Nanopositioning for Lithography and Data Storage

Wenbo Zhang, Ming Pang, and Changhai Ru

The history of the science shows that there is usually one strong social demand behind a new technology. The new technology, which can solve the key problem, must be got people's praise and attraction. Nanotechnology is also facing the similar historical opportunity. Nanotechnology, which is understanding matter and the control of matter at dimensions of 100 nm or less, appeared in recent years [1–4]. The inventions of the scanning tunneling microscope (STM) [5] and the atomic force microscope (AFM) [6], which are widely used in biology [7, 8], chemistry [9], materials science [10, 11], and physics [12], greatly promoted the research of the nanotechnology areas. The wide applications all depend on the high-precision nanopositioning stages. Based on this, nanopositioning is one of the key requirements of nanotechnology [13]. Nanopositioning involves precision control and manipulation of devices and materials at a nanoscale. Nanopositioners are precision mechatronic systems applied to move objects over a small range with a resolution down to a fraction of an atomic diameter. Hence, nanopositioners need to possess high resolution, accuracy, stability, and fast response. To realize manipulation at the nanometer scale with SPM, nanopositioning systems are necessary [14, 15]. In SPMs, nanopositioning system is used to control the probe to scan over a sample surface. Accurate position sensor and feedback control system are key to successful nanopositioning.

Nanopositioning systems are vital in optical alignment systems [16, 17] and next-generation space telescopes [18, 19]. Such nanopositioning is also crucial

W. Zhang • M. Pang (✉)
College of Automation, Harbin Engineering University, Harbin 150001, China
e-mail: guanglongriyue@163.com

C. Ru
College of Automation, Harbin Engineering University, Harbin 150001, China

Research Center of Robotics and Micro System & Collaborative Innovation Center of Suzhou NanoScience and Technology, Soochow University, Suzhou 215021, China

C. Ru et al. (eds.), *Nanopositioning Technologies*,
DOI 10.1007/978-3-319-23853-1_12

for biology areas, such as cell tracking and nanomaterials testing [20, 21], DNA analysis, the manufacturing of small objects [22], and nanoassembly [23]. Nanopositioning systems are the core of lithography tools, which might replace traditional optical lithography systems [24, 25]. The new lithography technology can construct semiconductor integrated circuits (ICs) having less than 35 nm nodes [26]. Similarly, the servo system of hard-disk drives (HDD) also requires nanometer-scale precision [26, 27]. The nanopositioning system can be used to realize the areal densities higher than 1 Tb/in^2 in probe-storage device [28, 29]. Novel nanopositioning technology is also required for semiconductor inspection [30, 31]. The vast range of applications necessitates nanopositioning systems with high resolution, high bandwidth, and robust control designs [32–38].

This chapter introduces some typical applications of nanopositioning, including lithography, data storage, and fault analysis.

12.1 Nanopositioning for Lithography

Lithography technology is the key of the semiconductor industry. Improvement of lithography has been a key driver for improving integrated circuit (IC) performance. Therefore, new generation lithography systems based on high precision nanopositioning and nanometer precise placement have attracted much attention [39]. At the same time, low-cost technologies for mass producing nanometric structures are in dire demand, especially for future manufacturing of integrated optics [40]. Diffraction grating is an important dispersion element in modern physics [41], and the diffraction efficiency of grating has strict requirements on the notch groove depth and shape [42, 43]. Mechanical characterization of chips is the best way to make diffraction grating by ruling engine. The requirements of ruling engine are large useful area, high precision and multifunction. Large useful area means large ruling width and ruling depth. High precision means the positional accuracy of optical grating. Multifunction means multiple inputs get different ruling results.

Electron beam, which is called electron projection lithography (EPL) [44, 45], is used for patterning the substrate to replace traditional optical lithography. Scanning electron beam can realize 10 nm lithography [46, 47]. Fischer and Chou used a modified scanning electron microscope to realize 10 nm electron beam lithography and sub-50 nm overlay [47]. They repeatedly achieved gratings of 10 nm wide on thick GaAs substrates, as shown in Fig. 12.1.

The disadvantage of scanning electron microscope is uneconomical and has low throughput due to its inherent drawback. X-ray lithography can realize high throughput with 50 nm resolution [48]. However, X-ray lithography tools are rather expensive. Chou et al. developed nanoimprint lithography (NIL), which is a low-cost mass manufacturing technology, for fabricating sub-25 nm features [40]. The fabrication of the mold used for NIL process is one of the most important research hotspots of NIL.

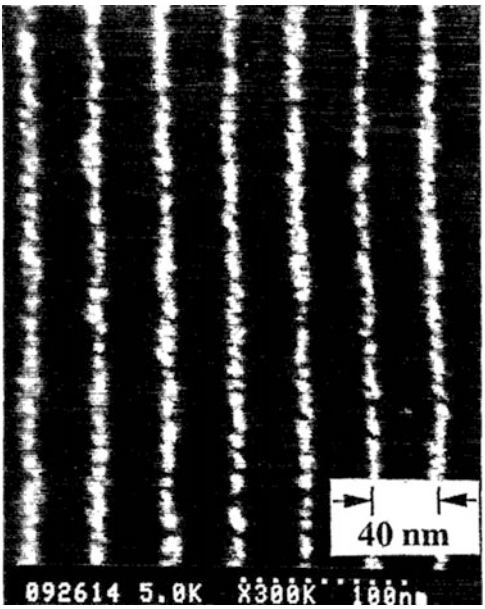


Fig. 12.1 Scanning electron micrograph of a 40 nm period Ni/Au grating with 10 nm linewidth on bulk GaAs [47]. Copyright 1993. Reproduced with permission from Applied physics letters

Vorbringer-Doroshovets et al. demonstrated nano-scale lithography by using a scanning probe combined with a Nanopositioning and Nanomeasuring Machine [49], as shown in Fig. 12.2.

For the scanning probe lithography experiments, the researchers used a modified hybrid Atom Force Microscope (AFM) set-up to operate at ambient conditions and room-temperature. The results are shown in Fig. 12.3.

For the application of lithography in grating ruling machine, Harrison successfully applied interference to the notch groove positional servo system and developed MIT-A [50] and MIT-B [51]. For MIT-A, the notch groove position is confirmed based on the optical wavelength. The area and precision of the optical grating were improved. Thereafter, grating interference measuring technique, laser interference measuring technique, piezoceramic micro-positioning technology, and computer control technology were developed and applied to optical grating. In 1966, Harrison et al. developed MIT-C [52] based on the experience of MIT-A and MIT-B. This grating ruling machine was constructed out of the forth all-purpose measuring machine produced by Moore Company, as shown in Fig. 12.4.

Compared with traditional grating ruling machine, the measuring elements for feedback were implemented by the single frequency laser interferometer and the feedback actuator was implemented by the differential gear train. MIT-C has double work stage structure, which is connected by flexure hinge, the inner stage is used to correct the straightness error of the guide rail. Now MIT-C has ruled the optical grating, of which the max size is 340×440 mm.

Since Harrison and Stroke successfully used the interference servo system in the grating ruling machine, many researchers developed various kinds of ruling

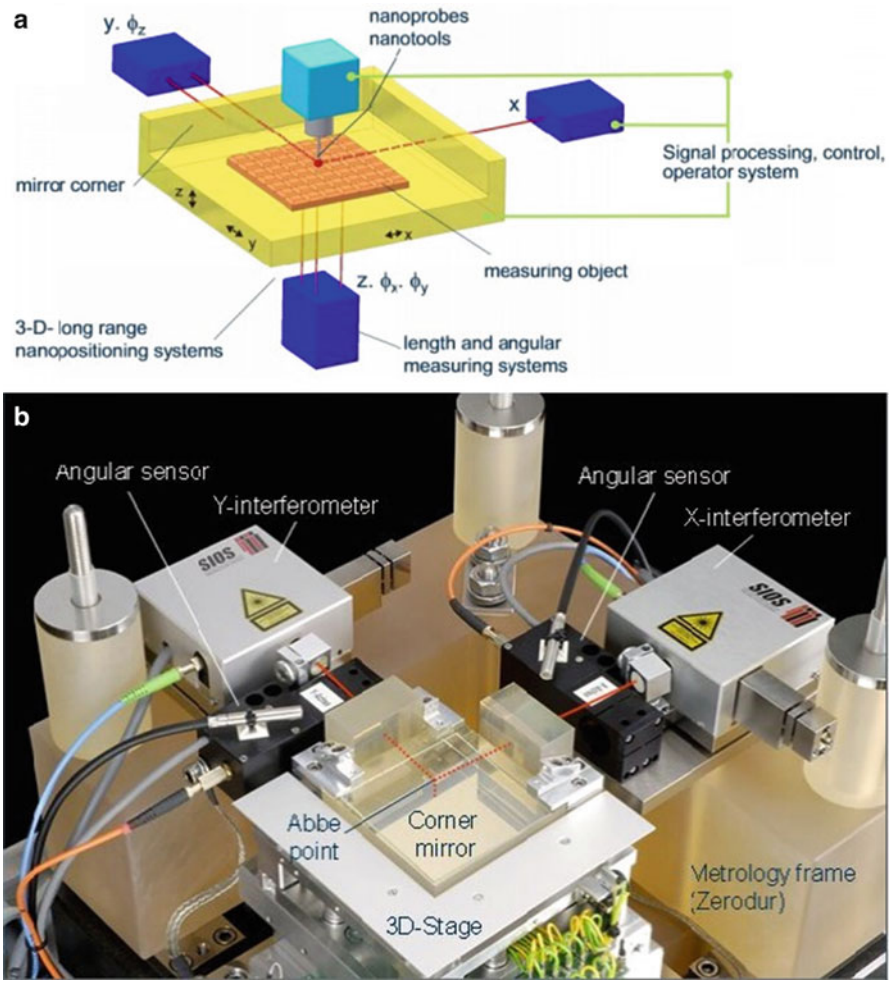


Fig. 12.2 (**a**) Basic concept and main components of the Nanopositioning and Nanomeasuring Machine (NPM Machine). (**b**) Additional angular sensors for measurement [49]. Copyright 2013. Reproduced with permission from SPIE Advanced Lithography

machines. Stiff et al. developed interference servocontrol grating ruling engine [53]. In 1992, Kita and Harada in the Central Lab of Hitachi developed digital control grating ruling engine [54], as shown in Fig. 12.5.

It has both macro- and micro-positioning systems for the first time. The macro-positioning system is driven by the screw nut structure and the micro-positioning system with upper and lower layers of platform structure is driven by piezoelectric ceramic. Its pivot swinging angle of workbench in 300 mm stroke is 0.2. The maximum grating groove spacing is 10,000 g/mm and the dividing control precision is 5 nm [54].

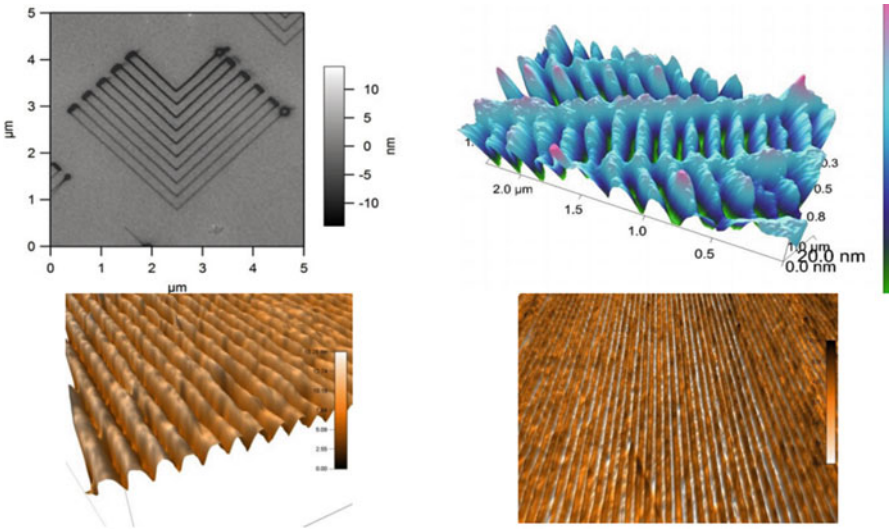


Fig. 12.3 AC-AFM mode images of the SPL-structures generated in calixarene resist. All lithographic steps and AFM-images were done step by step with the same cantilever (and without any development steps). The two bottom AFM topography images show line pattern written with 45 nm half-pitch into 20 nm thick C-MC4R resist [49]. Copyright 2013. Reproduced with permission from SPIE Advanced Lithography

Fig. 12.4 MIT-C grating ruling machine

The control strategy is not only decided by the drive method but also the controlled member. There are mainly two control strategies of nanopositioning for lithography, including feedback control strategy and feedforward control strategy. PID Control is a classical feedback control method. The integral controller of PID

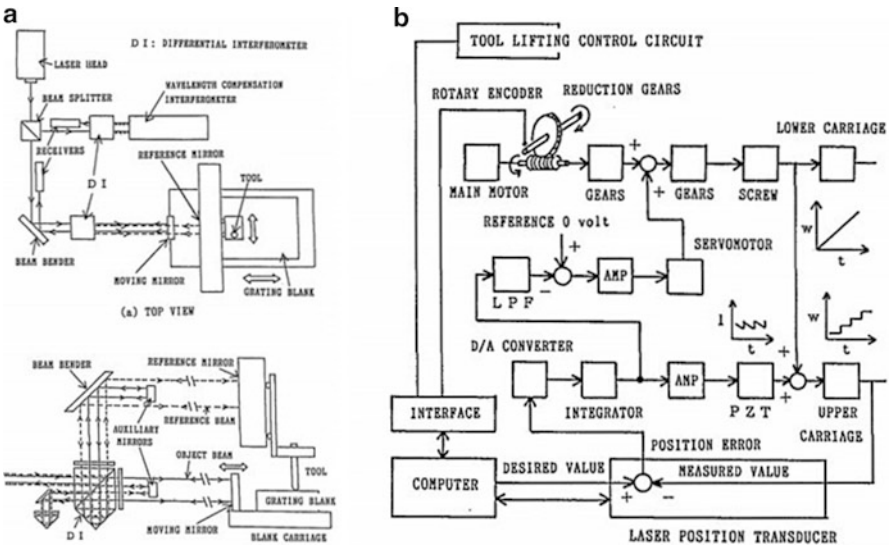


Fig. 12.5 Hitachi Grating Ruling Engine. (**a**) Optical layout of the laser interferometer. (**b**) Control system block diagram: *LPF* low-path filter, *PZT* piezoelectric device, *AMP* amplifier [54]. Copyright 1992. Reproduced with permission from OSA Publishing

can provide high gain feedback in low-frequency stage, which can be used to overcome creep and hysteresis. It is suitable to be used in nanopositioning systems. Kouno et al. and Cuttino et al. used traditional PID feedback controller and dynamic compensation module to reduce the hysteresis-caused error of the piezoelectric actuators in the single point diamond lathes less than 150 nm [55, 56] (Fig. 12.6).

Tan et al. designed a learning controller, which can auto-control the parameters of PID on line, which improved the robustness of the integral controller [57]. Heertjes et al. proposed a nonlinear PID controller, which can modify the parameters of PID on line [58]. Compared with traditional PID controller, this nonlinear PID controller can suppress interference and improve positioning accuracy, which was proved by positioning on the wafer table. Maeda et al. designed a controller suitable for super precision machine tool with traditional PID, feedforward and band-pass filter, which reached a high positioning accuracy to track a circle at 500 Hz [59] (Fig. 12.7).

The disadvantage of PID is that it needs higher feedback gain to realize positioning accuracy on nanometer in high frequency band while piezoelectric actuator can only offer low gain at high frequencies. Numasato and Tomizuka used Notch filter between the two grades drivers of HDD, which are voice coil motor and piezoelectric ceramics, to solve the problem. The Notch filter can alter the first harmonic peak of the control system and improve control precision [60] (Fig. 12.8).

Leang and Devasia added Notch filter to the control system of Atomic Force Microscopy, which improved the gain margin of piezoelectric actuators and suppressed the disadvantages of piezoelectric actuation, such as hysteresis, creep, and vibration [61], as shown in Fig. 12.9.

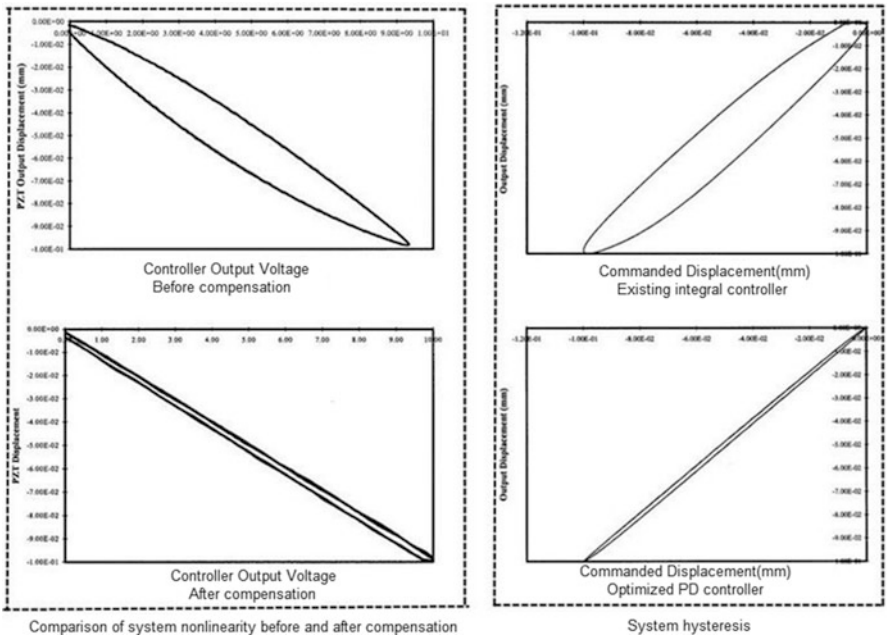


Fig. 12.6 Experimental results [56]

It can be seen that closed-loop control is widely used in nanopositioning for lithography, but it is limited by parameter uncertainty and high-frequency interference. Hence, other control methods such as adaptive control androbust control were considered. Cruz-Hernández and Haywar used lead-lag control to compensate the lag in the energy transfer process of smart materials [62]. Shakir and Kim used lead lag controller and LQG controller to realize a resolution of 5 nm [63]. The result is shown in Fig. 12.10.

Schitter et al. proposed a control strategy based on H_∞ theory, which is used in AFM (Atom Force Microscopy) control system and making the scanning speed five times higher than traditional PID control while the error was also reduced [35]. For a given scan rate, the controller error of the H_∞ controller was always lower than that of the PID controller. The average error (measured at the edges of patch) of the H_∞ controller was from one-fifth to one-third of the PID controller at 20 and 4 Hz, respectively. Furthermore, H_∞-controlled AFM had better damping of the stick slip effect. The experimental results are shown in Fig. 12.11.

Helfrich et al. combined H_∞ control with iterative learning control based on feedforward and obtained relatively good positioning effect [64]. Bashash and Jalili used robust adaptive control method on coupled parallel piezoelectric bending nanopositioning stage to suppress the lag of piezoelectric actuator and coupled nonlinearities. The positioning accuracy of the system was significantly improved [65]. Sebastian and Salapaka designed an H_∞ robust control method

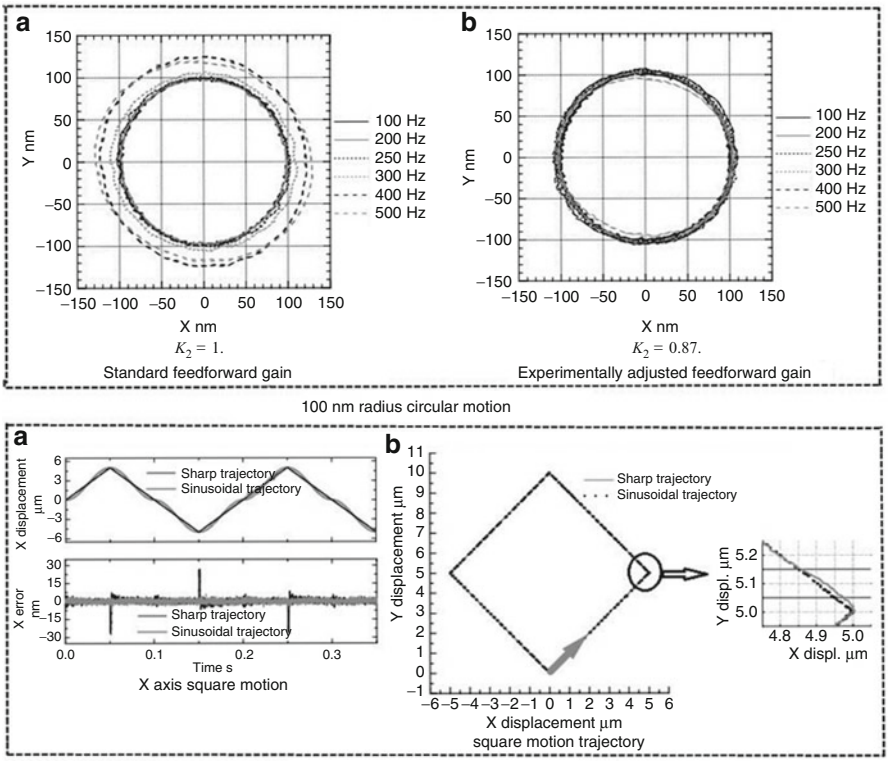


Fig. 12.7 100 nm radius circular motion and square motion in the micrometer range [59]

for AFM nanopositioning. The control system not only eliminated the influence of nonlinearity but also much improved the positional accuracy and speed [66]. In addition, to use the model information better, many nanopositioning systems used feedforward control or combined feedforward control and feedback control in the controlling process. Kuhnen and Janocha used feedforward control method to reduce nonlinear errors from 30 to 3 % [67]. Figure 12.12 shows the compensated characteristics of the overall system given by the serial combination of the inverse feedforward controller and the magnetostrictive actuator.

The one-dimensional ultra-precision stage developed by Tokyo Institute of Technology, which is controlled by the combination of conventional PID and feedforward compensation and driven by PMLSML (Permanent Magnet Linear Synchronous Motor), realized 2 nm step positioning [68]. Yong et al. used feedforward control on the *X*-*Y* nanopositioning stage of a scanning probe microscopy, and achieved maximum error of 2 % [69].

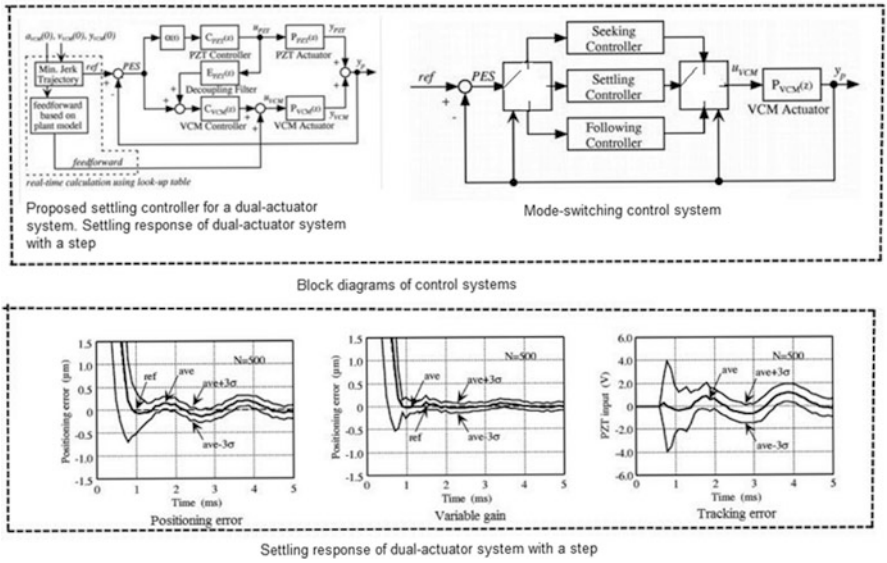


Fig. 12.8 Block diagrams of control systems and settling response of dual-actuator system with a step [60]

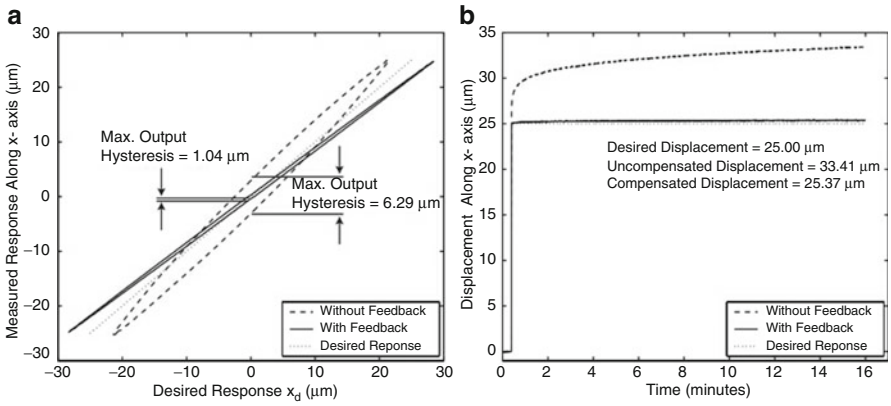


Fig. 12.9 Experimental results showing (**a**) hysteresis and (**b**) creep compensation using high-gain feedback. *Dashed line* is without compensation; *solid line* is with compensation; and *dotted line* is the desired response [61]. Copyright 2002. Reproduced with permission from the POL publication

12.2 Nanopositioning for Data Storage

There is a great need to develop ultra-high density, small form factor, and low cost memory devices since these data-storage technologies can bring enormous impact to the areas of computer science and mobile handheld devices. Due to the limits of

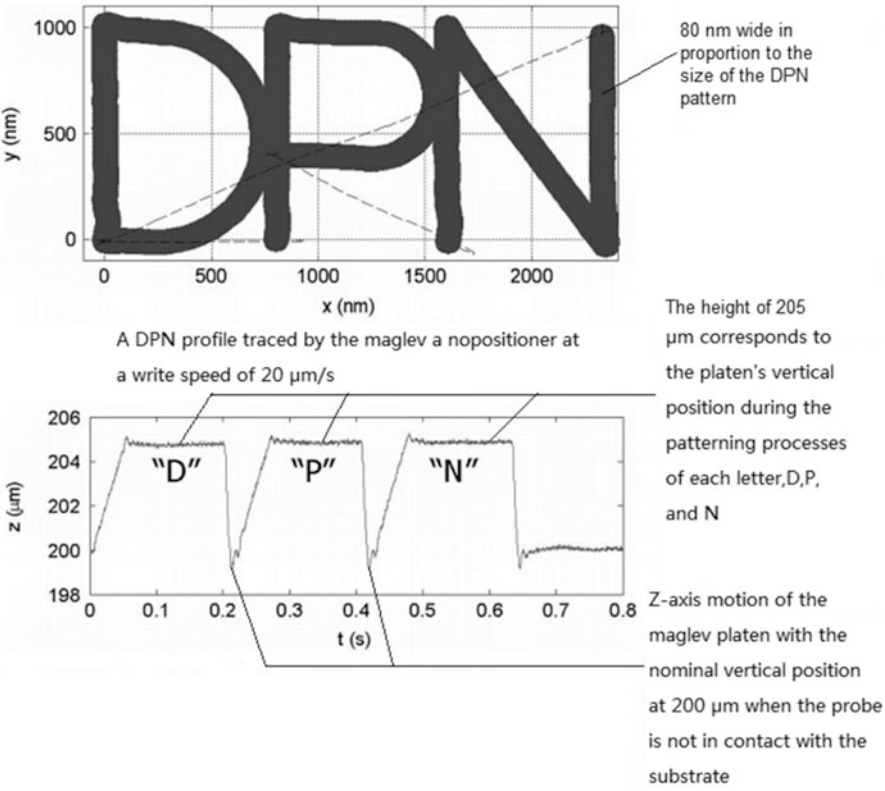


Fig. 12.10 Experimental results [63]

physics, the scale and storage capacity of conventional storage technologies, such as magnetic recording and solid-state flash memory, are reaching their limits. New technologies are proposed to address future data-storage requirements.

High capacity, small-form-factor, and low-power are the primary advantages of probe storage. With a sharp tip scanned in close proximity to a surface, the STM and AFM have the ability to image or modify surfaces with spatial resolution ranging from angstroms to tens of nanometers [70, 71]. STM and AFM have potential for data-storage device and can realize higher areal data storage densities than conventional methods [72]. In AFM, a sharp tip is fixed on the end of a soft cantilever spring to sense and apply atomic force [6]. AFM uses high-frequency, piezoresistive silicon cantilevers to mechanically read back topographic data (Fig. 12.13).

The example is shown in Fig. 12.14, where individual xenon atoms were positioned to spell out "IBM."

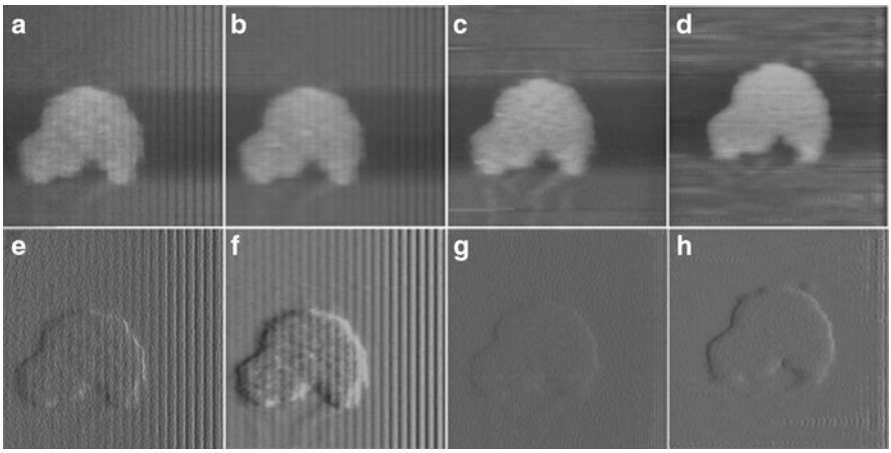


Fig. 12.11 Purple membrane imaged with an H_∞-controlled AFM [(**a**, **c**, **e**, and **g**)] and a PID-controlled AFM [(**b**, **d**, **f**, and **h**)]. All images were recorded from right to left and are 1.5 by 1.5 μm^2. The line scan rate is 20 Hz for (**a**, **b**, **e**, and **f**) and 4 Hz for (**c**, **d**, **g**, and **h**). The sample topography is shown in (**a**–**d**) [scaling from *black* to *white* is 23 nm for (**a**, **c**) and 25 nm for (**b**, **d**)]. The cantilever's deflection signal is shown in (**e**, **h**) [scaling is 4 nm from *black* to *white*] [35]. Copyright 2001. Reproduced with permission from AIP publisher

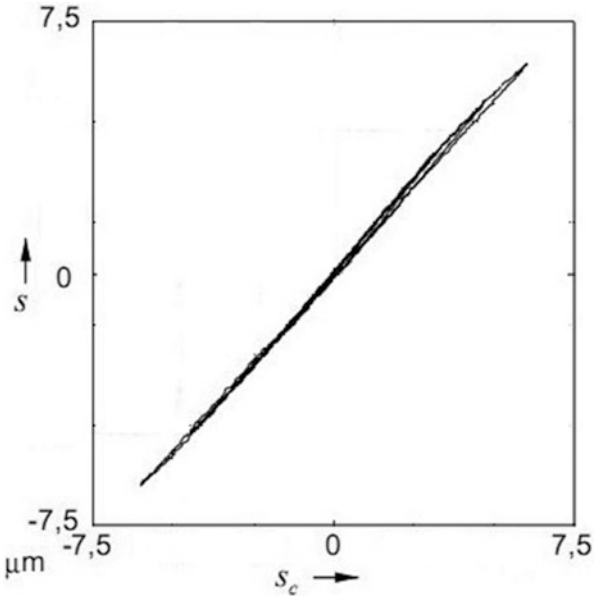


Fig. 12.12 Compensated hysteretic displacement-given displacement relation [67]

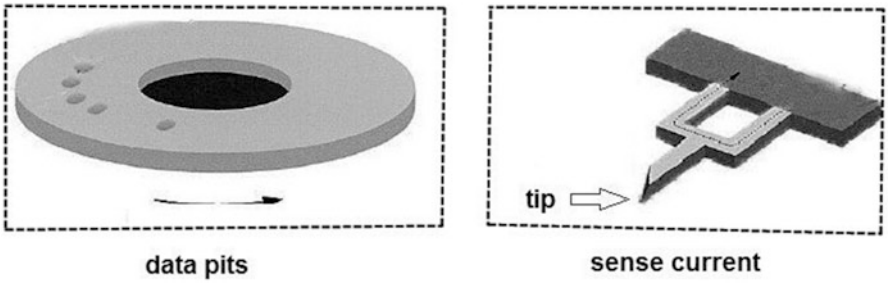


Fig. 12.13 Concept of topographic data storage using an AFM tip. The tip rides over the surface of the substrate, causing deflection of the cantilever as the tip moves over a topographic feature. The deflection is detected via a piezoresistive sensor [72]

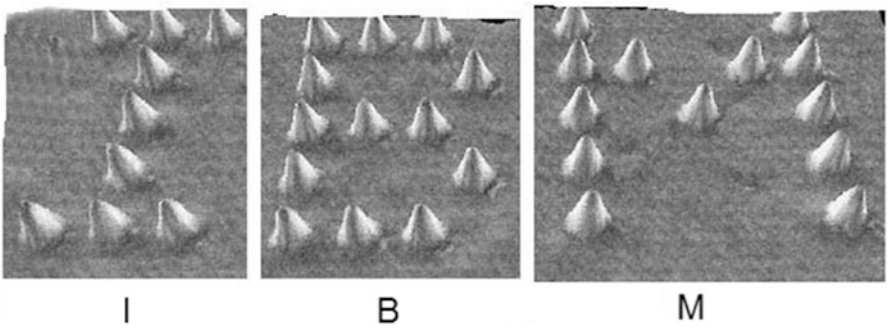


Fig. 12.14 Atomic-scale message demonstrating the ability to position single atoms with angstom precision. Each bump is an individual xenon atom positioned with the tip of a scanning tunneling microscope operating at 4.2 K. This represents perhaps the ultimate in data-storage density, with a single atomic "bit" (courtesy of D. Eigler, IBM Research Division, Almaden Research Center) [72]

STM scans the tip while monitoring the tunnel and maintaining tip position servo current constant to map out the surface topography [73]. The actuator can move in the vertical and cross-track directions. A closed-loop servo system, of which the bandwidth is nearly 1000 Hz, used the signal from the piezoresistive cantilever as the error signal to control the vertical load on the cantilever [74].

However, the limitation of the feedback speed and low tunneling currents makes the atomic-level storage based on AFM difficult to realize. The micro-electro-mechanical systems (MEMS)-based arrays of cantilevers operating in parallel was employed to as a solution, with each cantilever performing WRITE/READ/ERASE operations in an individual storage field. In MEMS-actuated magnetic probe-based storage system, a magnetic storage medium is positioned in the x/y plane [75]. Writing is achieved by an array of probe tips, each tip was positioned by a z-axis positioning system. The position servo can realize displacements up to 55 % of the initial actuation gap.

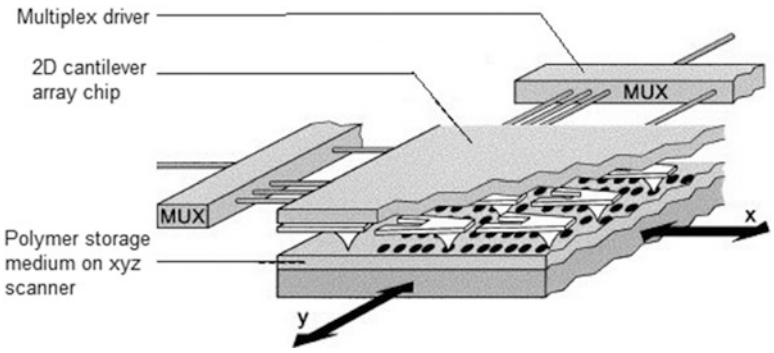


Fig. 12.15 Illustration of the Millipede concept [77]

Vettiger and Eleftheriou presented the scanning-probe-based data-storage concept, which is called the "millipede" [76, 77]. The technology uses the thermomechanical concept to record and recover data in thin polymer films, as shown in Fig. 12.15.

It is encoded as nanometer-scale indentations, for which the presence or absence of indentations corresponds to logical ones or zeros, respectively. Early experimental results demonstrated that the overall chip tip-apex height control was within 500 nm. It is realized by using feedback control [78]. The main drawback of thermomechanical probe-based data is the low data rate per microcantilever. To solve this, WRITE/READ operations depend on a mechanical parallel *x*/*y* scanning of either the entire cantilever array chip or the storage medium [77]. A row/column time-multiplexed addressing scheme is proposed to ensure the parallel operations efficient. It addresses the array column by column with full parallel WRITE/READ operation within one column [76]. The cantilever moves from one bit position to the next and is driven by the pulse. During reading process, each tip is always positioned over the center of a desired track. Similarly during writing, the tips should be positioned on the written marks, which are aligned in a predefined way. The servo loop for track following is driven by an appropriate error signal, which is called PES. The stability and robustness of the associated tracking servo loop is directly affected by the quality of the PES [79]. Magnetic recording provides the concept of mutually vertically displaced to provide two signals in quadrature, which can be combined to provide a robust PES [79]. As shown in Fig. 12.16a, servo marks are placed in bursts labeled A and B for the in-phase signal and C and D for the quadrature signal. The tip moves from the edge of the top mark toward its center, then toward its bottom. The combination of the two signals (I and Q), shown as solid lines in Fig. 12.16b, has zeros crossing as all track center locations and constant slope, which qualifies it as a valid PES. At even-numbered tracks, the zero of the in-phase signal indicates the track center. In turn, the zeros of the quadrature signal can be uniquely mapped into a position estimate by examining the polarity of the in-phase signal at corresponding positions.

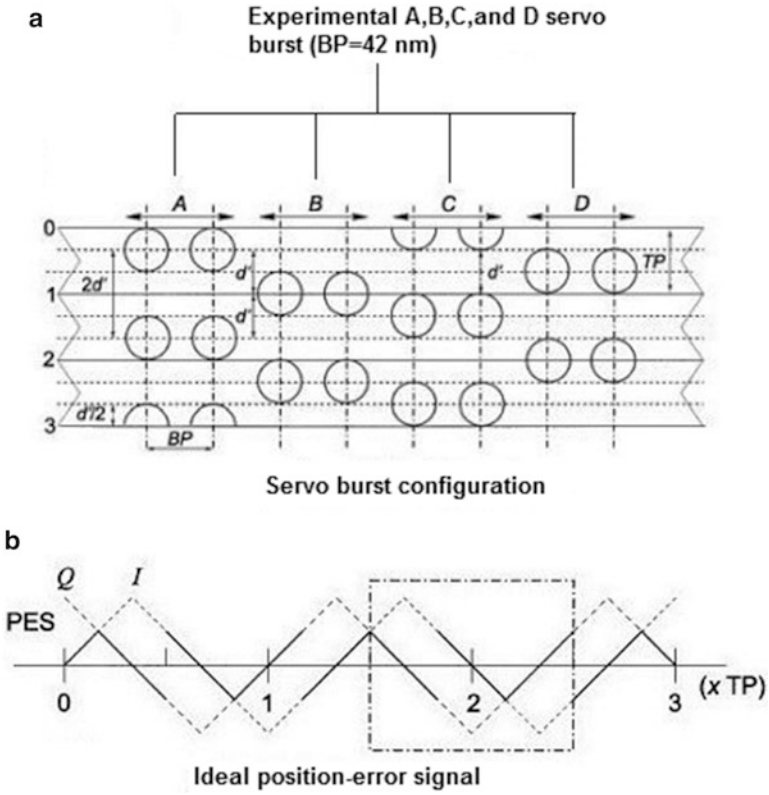


Fig. 12.16 Only two signal marks per burst are shown here to simplify the presentation while in fact, each burst typically consists of many marks in order to enable averaging of the corresponding readout signals. The *solid horizontal lines* depict the track centerlines and *circles* represent written marks [77]

The experimental results of the principle of PES generation based on servo and the resulting in-phase and quadrature signals are shown in Fig. 12.17.

Experiments with single cantilevers demonstrated that data can be recorded at a density of 641 Gb/in^2 and read back with raw error rates better than 10^{-4} when the *x*/*y* actuation is performed by an off-the-shelf piezoscanner with closed-loop position control via internal capacity sensors [80]. The error-rate performance as a function of off-track position with 641 Gbit/in^2 is showed in Fig. 12.18.

The displacement control problem of microscanner can be divided into *X* and *Y* directions, with low interacting. So the microscanner design can be used for parallel-probe-based mobile storage devices or other nanopositioning system [81]. Because the displacement of the scanner makes a change in the heater temperature, two pairs of thermal position sensors are fixed on the cantilever-array chip to sense a displacement of the object by temperature, as shown in Fig. 12.19.

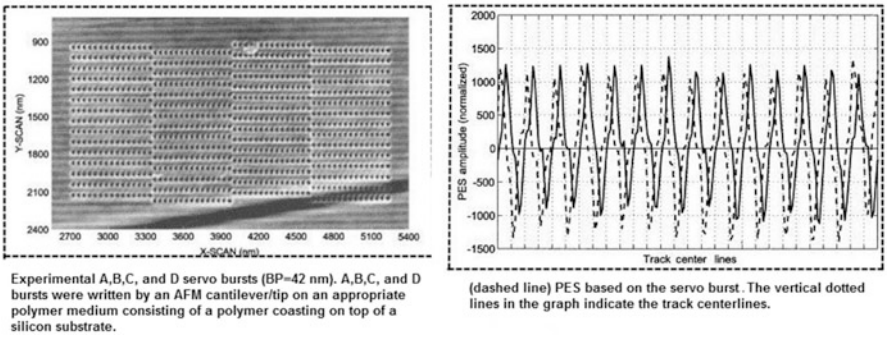


Fig. 12.17 Experimental results [77]

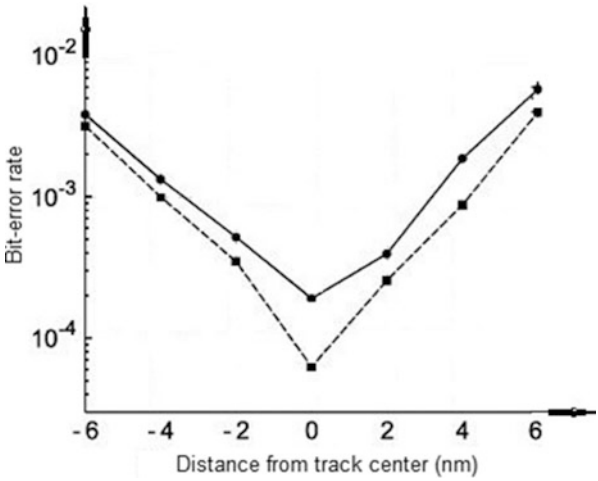


Fig. 12.18 Raw bit-error rate as a function of distance from the track centerline [80]

A current-to-voltage amplifier is used to measure the changes in the resulting current. The sensors are used in pairs using a differential configuration to minimize drift effects. The experiments result that the linearity, drift, and noise limit perform good, even the device is simple. These sensors' bandwidth can reach 4 kHz and the displacement resolution is less than 1 nm [83].

Because the fabrication of the electrostatic microactuators is easy, the electrostatic microactuators are widely used in the MEMS field and data storage. A complete electrostatic microactuators has two parts. The stator is fixed while the shuttle is movable. When voltage is applied, the electrical force is generated between the two combs. The overlap area of the fingers and the distance between the electrodes increase, as shown in Fig. 12.20.

Engelen et al. design a prototype *x*/*y* scanner using electrostatic comb drive actuation, as shown in Fig. 12.21.

In the *z* direction, etching the device through the full thickness of the wafer (400 μm) leads to a large out-of-plane stiffness for passive shock. The large etch

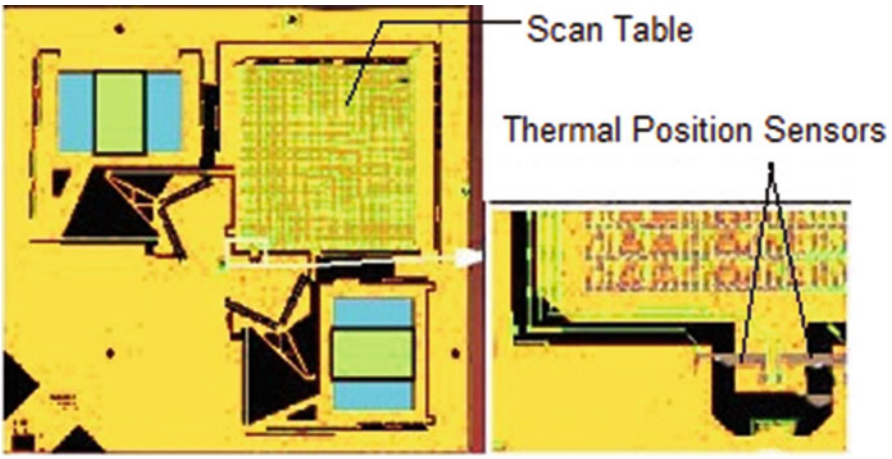


Fig. 12.19 Microscanner and thermal position sensors [82]

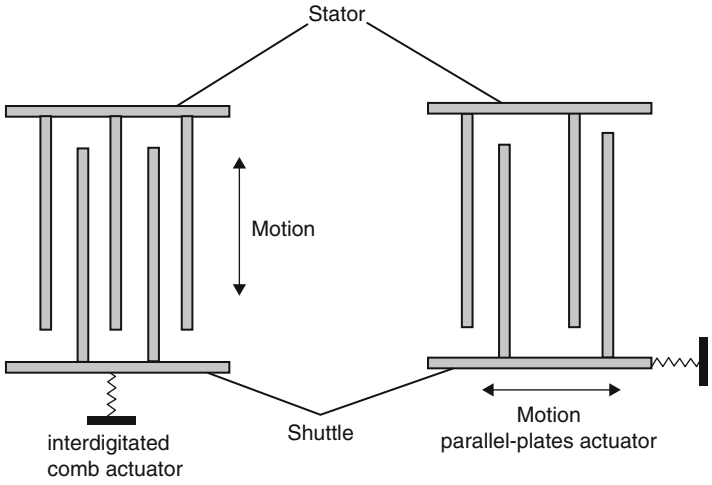


Fig. 12.20 Electrostatic micro-actuators [13]

depth is necessary for the use of a relatively large gap between fingers (25 μm) but it reduced the actuation force. To increase the force without decreasing the minimum etch trench width, a taped shape [85, 86] and a new 'stepped' shape, shown in Fig. 12.22.

The simulation result of the three finger shapes is shown in Fig. 12.23.

From the simulation result it can be got that stepped shape is better than the other two. To provide the margin for control operations, the range of achievable displacement needs to be increased by increasing the force available at large positive displacements. The drawback of the interdigitated comb drives is that it generates weak force while it needs high voltage.

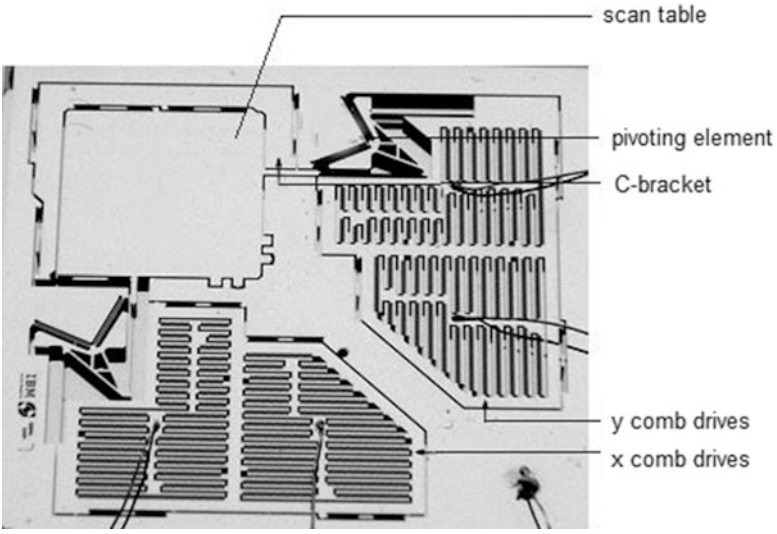


Fig. 12.21 Photograph of a fabricated device (2 × 2 cm). The wires are electrical connections to the isolated stators and to ground the moving parts of the device (*bottom right*) [84]

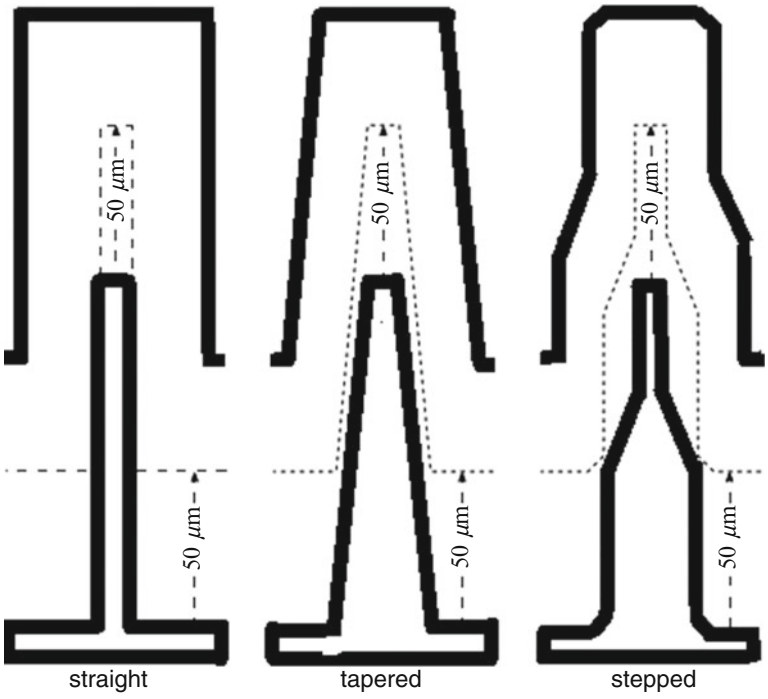


Fig. 12.22 The three investigated finger shapes, showing the new stepped finger shape on the right [84]

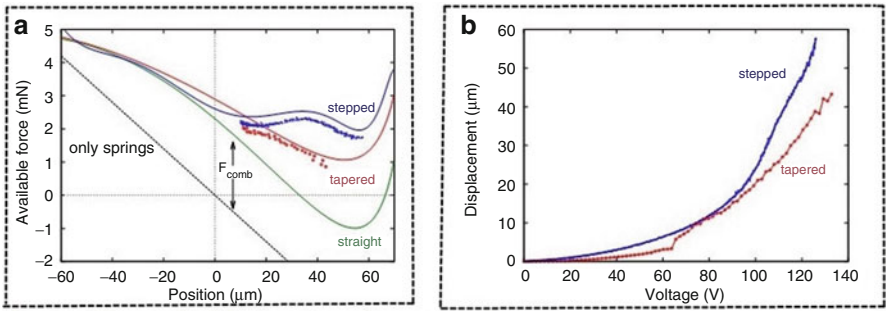


Fig. 12.23 Simulation results [84]

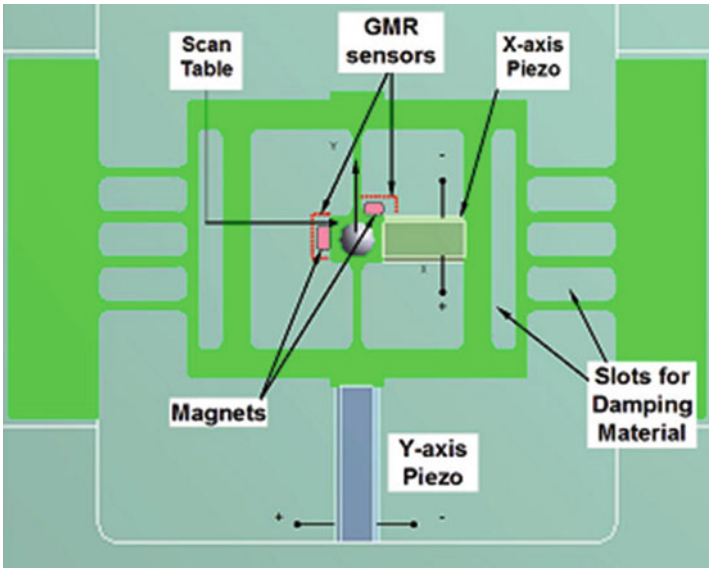


Fig. 12.24 Schematic depiction of the *X*-*Y* scanner with the MR-based position sensors [87]. Copyright 2010. Reproduced with permission from the POL publication

Kartik et al. develop a high-speed *x*/*y*-positioning scanner design with high-bandwidth piezoelectric actuator [87], as shown in Fig. 12.24. A suitable high-speed scanner for nanopositioning system should have high first resonances and a low coupling between the three orthogonal directions.

The high first resonances are realized by the high stiffness of the flexures and low carrier. The low cross-coupling is achieved by low mass and uncomplicated mechanical connections. The scanner comprises one stage each in the *X*- and *Y*-scan directions, which is actuated by piezoelectric stacks. Piezo-electric stack actuators have high stiffness and ability to generate large actuation forces, which is

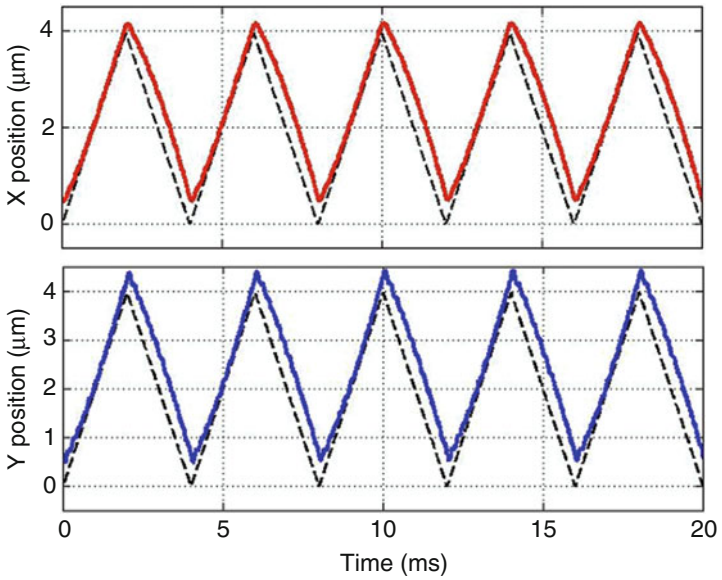


Fig. 12.25 The response of the scanner along the *X*- and the *Y*-scan directions to a 250 Hz triangular drive signal [87]. Copyright 2010. Reproduced with permission from the POL publication

suitable for such application. Due to the requirement for high scan speed and high actuation bandwidth, the allowable capacitances of the selected piezostacks have an upperbound. The maximum achievable elongation and blocking force of the piezo-actuator is also constrained. The effective mass carried by the *X*-*Y* scan table has to be reduced in view of the limit imposed by the requirement for speed on the flexural rigidity of the scanner. So in the *Z*-direction, the scan table is not actuated. Piezo-actuators have hysteresis characteristic, which has bad effect on the scanner. When a triangular drive signal at a frequency of 250 Hz applied at either of the *X*- and *Y*-directions, the displacement of the scan table is shown in Fig. 12.25.

From the trace of Fig. 12.25, it is apparent to see the presence of hysteresis in the displacement response relative to the reference. It is a behavior typical of piezo-actuators when driven with a voltage amplifier. To solve this, Fleming et al. place charge amplifiers as the driving source instead of voltage amplifiers [88]. The positional signal is provided by Magneto-Resistance (MR) sensors. MR sensors are traditionally used as magnetic sensing [89]. The external magnetic field changes the electrical resistance of spintronics materials. The magneto-resistance effect is caused by the scattering of electrons which is measured by their mean free path length. Giant magneto-resistance (GMR) sensors, which is a frequently used advanced MR sensors, consist of ultra-thin layers and pinned magnetic layers. Daughton et al. have demonstrated that the MR sensors have high sensitivity, low noise, and high bandwidth [90], which achieve high resolution position sensing. The key problem is how to translate the motion of the scanner into a change in the

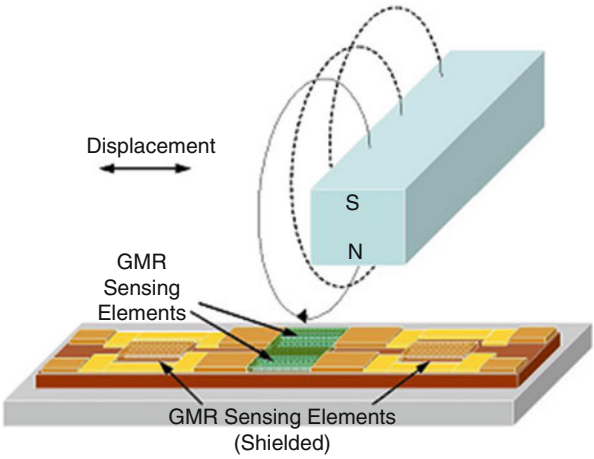


Fig. 12.26 Schematic depiction of the magneto-resistance based position sensing concept [87]. Copyright 2010. Reproduced with permission from the POL publication

magnetic field, which can be seen by an MR sensor. Sahoo et al. proposed a similar idea for topography sensing using micro-cantilevers [91]. The schematic depiction of the MR-based position sensor is shown in Fig. 12.26.

Four GMR sensors are configured into a Wheatstone bridge. Two of them are active GMR sensing elements while the other two are shielded GMR sensing elements. The sensor is sensitive in one direction in its plane with a cosine scale off in sensitivity as the sensor is rotated away from its sensing direction. The open-loop dynamics of the high-speed scanner were characterized by both the laser Doppler vibrometer and the MR-based position sensor. A laser Doppler vibrometer was used to calibrate the MR-based position sensors. When a white noise excitation is applied to the actuating piezo-stack in the corresponding direction, the open-loop frequency response of the stage in the X- and Y-directions is measured by the MR-based position sensor. The result is shown in Fig. 12.27.

The measured frequency responses for both axes can be fitted as sixth-order transfer functions, which can be used for controller design. A block diagram describing the closed loop is shown in Fig. 12.28.

The X- and Y-axes were controlled independently using H_∞ control framework, as shown in Fig. 12.29.

W_p and W_u first-order weighting functions to specify the closed loop transfer functions. W_p captures the requirement on tracking, and W_u ensures that the control effort is within limits and that the higher unmodelled resonance modes are not overly excited. The signal w_1 is known as the "exogenous input" and z_1 and z_2 are known as "exogenous outputs." The "tracking error" and "control signal" are denoted by v and u, respectively. The H_∞ control design problem is to find a stabilizing controller K_∞ such that $\|T_{zw}\|_\infty < \gamma_\infty$, where γ_∞ is a constant approximately equal to 1 [92]. Here

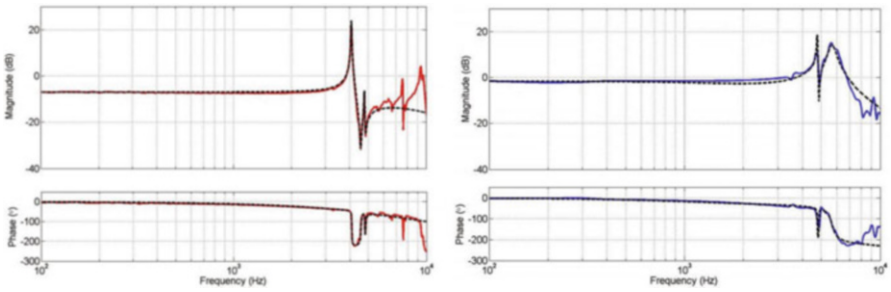


Fig. 12.27 *Left*: Open-loop frequency response of the scanner along the *X*-scan direction, measured with the MR sensor (*solid line*-type). *Right*: Open-loop frequency response of the scanner along the *Y*-scan direction, measured with the MR sensor (*solid line*-type); the corresponding sixth-order transfer function fit (*dashed line*-type) accurately captures the first two resonant modes [87]. Copyright 2010. Reproduced with permission from the POL publication

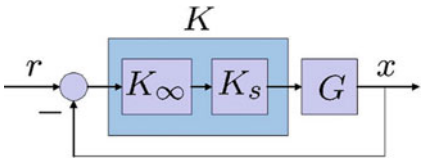


Fig. 12.28 Block diagram depiction of the closed loop architecture for each axis [87]. Copyright 2010. Reproduced with permission from the POL publication

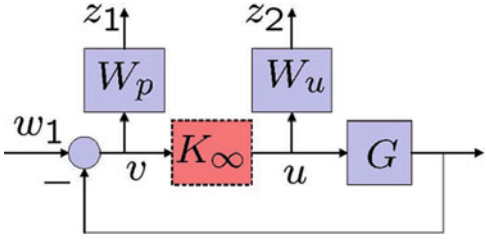


Fig. 12.29 Block diagram depiction of the H_∞ control formulation [87]. Copyright 2010. Reproduced with permission from the POL publication

T_{zw} is the closed-loop transfer function matrix relating z with w, where $z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$ and $w = [w_1]$. The H_∞ synthesis was used as the eighth order controllers, realized by a digital signal processor with a sampling period of 10 μs [87]. The experimental frequency response of the closed loop system for both the X- and Y-scan axes is shown in Fig. 12.30.

The closed loop system performance is shown in Fig. 12.31 when a triangular signal at 250 Hz, which is a typical reference for imaging applications, is applied along the *X*-scan direction, while stepping in the *Y*-scan direction by 100 nm. Since both the scan directions are controlled with very high bandwidths, minimal cross-coupling between the axes is seen.

Positioning for data storage system has a strict requirement for accuracy. Single position sensor is easily effected by drift and low-frequency noise. Pantazi et al. developed a medium-derived position-sensing concept which provides positional

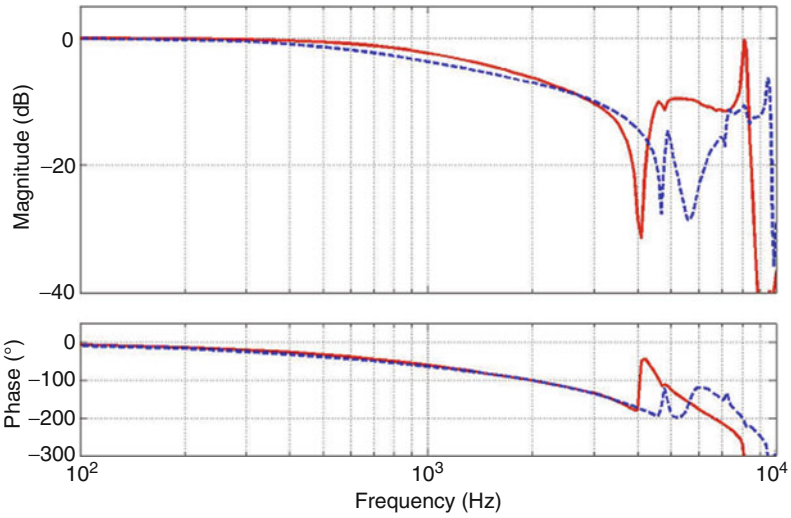


Fig. 12.30 Experimental closed loop frequency response for the *X*- (*solid line*-type) and *Y*-scan directions (*dashed line*-type) [87]. Copyright 2010. Reproduced with permission from the POL publication

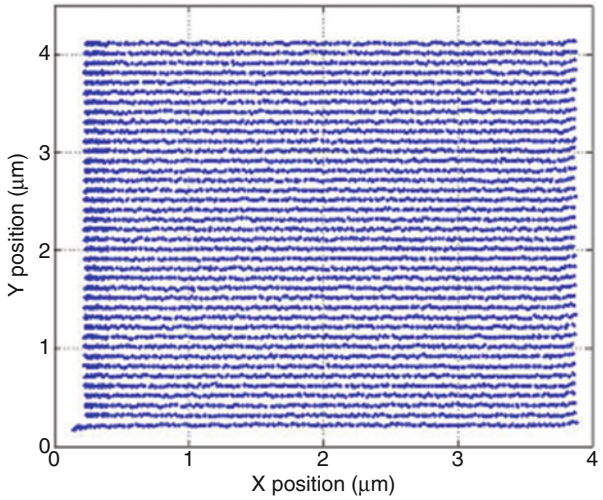


Fig. 12.31 A closed loop scan operation along the *X*-scan direction performed while stepping in the *Y*-scan direction by 100 nm [87]. Copyright 2010. Reproduced with permission from the 5th IFAC Symposium on Mechatronic Systems

information in the cross-track direction along the y-scan direction [93]. The "self-servo-write" process relies entirely on the thermal position sensors to move the scanner in closed-loop mode to the right position for inscribing the servo pattern [77]. Therefore, position accuracies are below the resolution of these sensors. Using

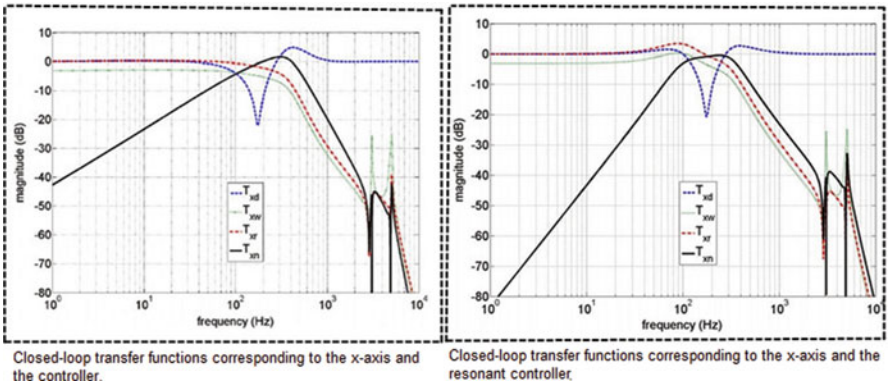


Fig. 12.32 *Left*: Closed-loop transfer functions corresponding to the *x*-axis and the controller. *Right*: Closed-loop transfer functions corresponding to the *x*-axis and the resonant controller [95]

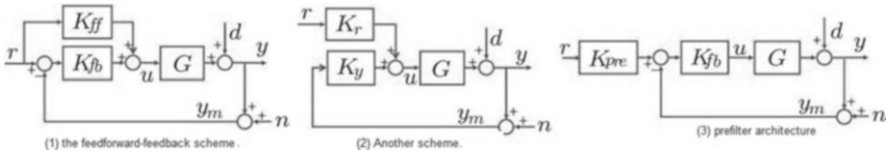


Fig. 12.33 2DOF control architectures: The schemes (*1*) and (*2*) are equivalent as control designs in that one can be retrieved exactly in terms of the other. Practical implementable designs for the controllers in (*1*) and (*2*) can easily be derived from the control design in (*3*), however the opposite situation may require certain factorization procedures [96]

this concept of directed shaping of the noise sensitivity transfer function, positioning accuracies can realize far below the noise levels of the position sensors without incurring a signification compromise on the tracking bandwidth [94]. When the scanner was only controlled in the frequency area where control was essential, the measurement noise doesn't impact the positioning accuracy in other frequencies [95]. So the noise was shaped. H_∞ control design approach and resonant controllers were implemented for the self-servo write process. The response of different signals is shown in Fig. 12.32.

Lee and Salapaka propose two-degree-of-freedom controllers, as shown in Fig. 12.33, to decouple the requirements for tracking performance and tolerance to measurement noise [96].

The drawback is that complexity of system increased in linear feedback control. Tuma et al. proposed a novel nonlinear control based on impulsive control to improve the tracking performance without compromise on bandwidth [97]. The principle of impulsive control is to change the controller states impulsively at discrete time instants, as shown in Fig. 12.34.

Using triangular waveform to evaluate this method, the simulation results of a range of controller gains and for a triangular reference signal of varying frequency were shown in Fig. 12.35.

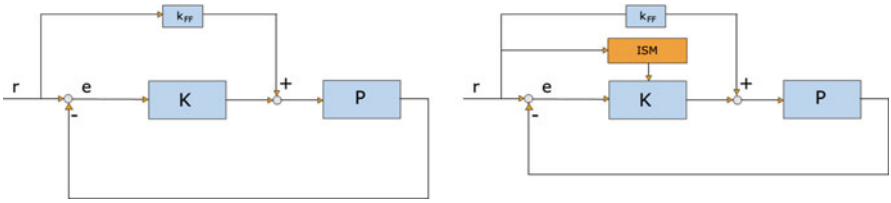


Fig. 12.34 *Left*: Linear feedback loop without impulsive state multiplication (ISM). *Right*: The same feedback loop with ISM [97]. Copyright 2010. Reproduced with permission from the POL publication

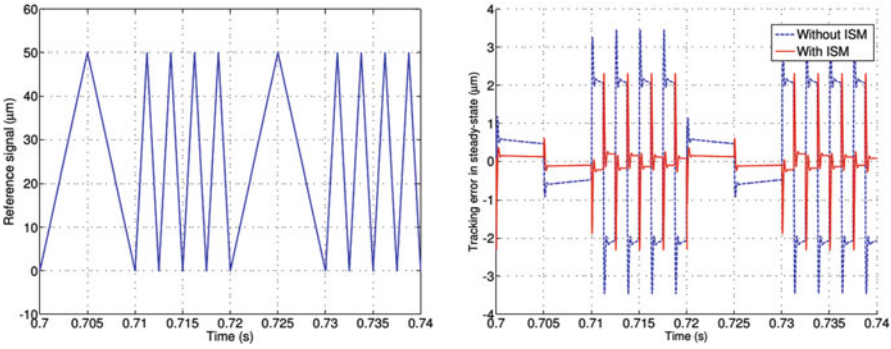


Fig. 12.35 *Left*: Reference signal. *Right*: Tracking error in steady state [97]. Copyright 2010. Reproduced with permission from the POL publication

For different controller gains k_i, the comparison of the tracking error for a triangular reference signal using the feedback loops with and without impulsive state multiplication (ISM) is shown in Fig. 12.36.

In the experiment presented, it can be shown that impulsive control offers better transient performance and easier implements. For piecewise affine reference signals, the stability conditions and performance bounds of impulsive control were derived [98]. Experiments indicate that in a high-speed AFM system, an impulse controller improved the tracking performance without increasing the bandwidth of the linear feedback loop [99].

The development of the nanotechnology has opened a new application area, especially in nanopositioning. This chapter has introduced some classical actuators, sensors, and control strategies applied in nanopositioning for lithography and data storage. It can be seen that nanopositioning must be applied in more area in the future.

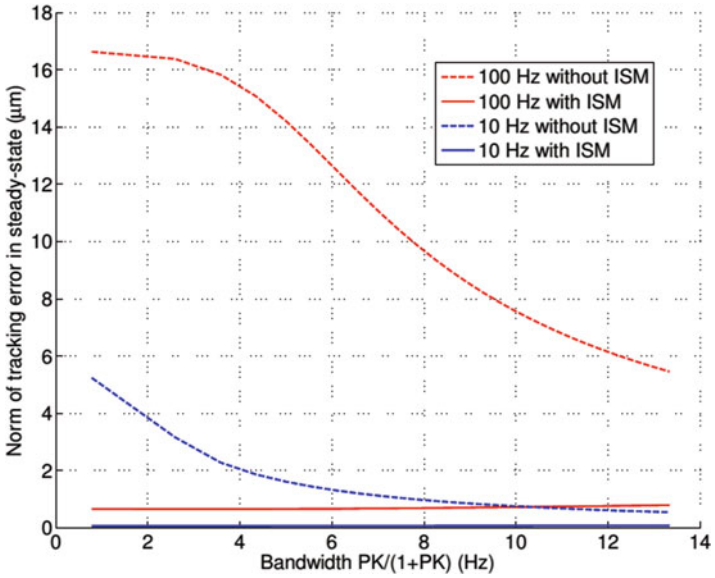


Fig. 12.36 The comparison of tracking error [97]. Copyright 2010. Reproduced with permission from the POL publication

Acknowledgement Thanks to the organizations and individuals for granting the pictures.

References

1. B. Shapiro, Workshop on control of micro-and nano-scale systems (2005)
2. T. Sheetz, J. Vidal, T.D. Pearson et al., Nanotechnology: awareness and societal concerns. Technol. Soc. **27**(3), 329–345 (2005)
3. Bhushan B. in *Springer Handbook of Nanotechnology*, ed. by B. Bhushan. Nanotechnology, vol 99 (Springer Heidelberg, 2004), p. 145
4. B. Bhushan, *Nanotechnology: Molecular Speculation on Global Abundance* (MIT Press, Cambridge, 1996)
5. G. Binnig, H. Rohrer, Scanning tunneling microscopy. IBM J. Res. Dev. **44**(1-2), 279–293 (2000)
6. G. Binnig, C.F. Quate, C. Gerber, Atomic force microscope. Phys. Rev. Lett. **56**(9), 930 (1986)
7. Q. Zou, K.K. Leang, E. Sadoun et al., Control issues in high-speed AFM for biological applications: collagen imaging example. Asian J. Control **6**(2), 164–178 (2004)
8. R. Kassies, K.O. Van der Werf, A. Lenferink et al., Combined AFM and confocal fluorescence microscope for applications in bio-nanotechnology. J. Microsc. **217**(1), 109–116 (2005)
9. K.D. Jandt, M. Finke, P. Cacciafesta, Aspects of the physical chemistry of polymers, biomaterials and mineralised tissues investigated with atomic force microscopy (AFM). Colloids Surf. B Biointerfaces **19**(4), 301–314 (2000)
10. I. Schmitz, M. Schreiner, G. Friedbacher et al., Phase imaging as an extension to tapping mode AFM for the identification of material properties on humidity-sensitive surfaces. Appl. Surf. Sci. **115**(2), 190–198 (1997)

11. K. Yamanaka, A. Noguchi, T. Tsuji et al., Quantitative material characterization by ultrasonic AFM. Surf. Interface Anal. **27**(5-6), 600–606 (1999)
12. M. Tsukada, N. Kobayashi, M. Brandbyge et al., Physics of artificial nano-structures on surfaces. Prog. Surf. Sci. **64**(3), 139–155 (2000)
13. S. Devasia, E. Eleftheriou, S.O.R. Moheimani, A survey of control issues in nanopositioning. IEEE Trans. Control Syst. Technol. **15**(5), 802–823 (2007)
14. M. Yves, *Scanning Probe Microscopes* (SPIE, Bellingham, 1995)
15. R. Wiesendanger, T. Mulvey, Scanning probe microscopy and spectroscopy. Meas. Sci. Technol. **6**(5), 600 (1995)
16. Z. Rihong, X. Daocai, Y. Zhixing, C. Jinbang, Research on systems for measurements of CCD parameters, in *Proceedings of SPIE Detectors, Focal Plane Arrays, Imaging Devices II* (1998), pp. 297–301
17. D. Krogmann, H.D. Tholl, P. Schreiber, et al., Image multiplexing system on the base of piezoelectrically driven silicon microlens arrays, in *3rd International Conference on Micro Opto Electro Mechanical Systems, MOEMS*, vol 99 (1999), pp. 178–185
18. J.C. Fasik. An inchworm actuator for the next generation space telescope (Burleigh Instruments, Inc., Fishers, 1998)
19. B. Zhang, Z. Zhu, Developing a linear piezomotor with nanometer resolution and high stiffness. Mechatron. IEEE/ASME Trans. **2**(1), 22–29 (1997)
20. Handbook of micro/nano tribology[M] (CRC, 1998)
21. S.V. Kalinin, D.A. Bonnell, Effect of phase transition on the surface potential of the BaTiO3 (100) surface by variable temperature scanning surface potential microscopy. J. Appl. Phys. **87**(8), 3950–3957 (2000)
22. D.R. Meldrum, W.H. Pence, S.E. Moody, et al., Automated, integrated modules for fluid handling, thermal cycling and purification of DNA samples for high throughput sequencing and analysis, in *Advanced Intelligent Mechatronics, 2001. Proceedings. 2001 IEEE/ASME International Conference on. IEEE* (2001), vol 2, pp. 1211–1219
23. G.M. Whitesides, L.J. Christopher, The art of building small. Sci. Am. **285**(3), 38–47 (2001)
24. Nanolithography: a borderland between STM, EB, IB, and X-ray lithographies[M] (Springer Science & Business Media, 2013)
25. P. Vettiger, U. Staufer, D.P. Kern, Preface. Microelectron. Eng. **32**(1), 1–2 (1996)
26. L.R. Harriott, Limits of lithography. Proc. IEEE **89**(3), 366–374 (2001)
27. R. Horowitz, T.L. Chen, K. Oldham, et al., in *Springer Handbook of Nanotechnology*, Microactuators for dual-stage servo systems in magnetic disk files (Springer Berlin Heidelberg, 2004), pp. 951–981
28. S. Hosaka, A. Kikukawa, H. Koyanagi et al., SPM-based data storage for ultrahigh density recording. Nanotechnology **8**(3A), A58 (1997)
29. A. Sebastian, A. Pantazi, G. Cherubini, et al., Nanopositioning for probe storage, in *American Control Conference, 2005. Proceedings of the 2005. IEEE* (2005), pp. 4181–4186
30. D.L. White, O.R. Wood II, Novel alignment system for imprint lithography. J. Vac. Sci. Technol. B **18**(6), 3552–3556 (2000)
31. S. Verma, W. Kim, H. Shakir, Multi-axis maglev nanopositioner for precision manufacturing and manipulation applications. Ind. Appl. IEEE Trans. **41**(5), 1159–1167 (2005)
32. A.J. Fleming, S.O.R. Moheimani, Precision current and charge amplifiers for driving highly capacitive piezoelectric loads. Electron. Lett. **39**(3), 282–284 (2003)
33. D. Croft, S. Devasia, Vibration compensation for high speed scanning tunneling microscopy. Rev. Sci. Instrum. **70**(12), 4600–4605 (1999)
34. D. Croft, G. Shed, S. Devasia, Creep, hysteresis, and vibration compensation for piezoactuators: atomic force microscopy application. J. Dyn. Syst. Meas. Control **123**(1), 35–43 (2001)
35. G. Schitter, P. Menold, H.F. Knapp et al., High performance feedback for fast scanning atomic force microscopes. Rev. Sci. Instrum. **72**(8), 3320–3327 (2001)
36. S. Salapaka, A. Sebastian, J.P. Cleveland et al., High bandwidth nano-positioner: a robust control approach. Rev. Sci. Instrum. **73**(9), 3232–3241 (2002)

37. O. M. El Rifai, K. Youcef-Toumi. Design and control of atomic force microscopes, in *American Control Conference, 2003. Proceedings of the 2003. IEEE* (2003) vol. 5, pp. 3714–3719
38. A. Stemmer, G. Schitter, J.M. Rieber et al., Control strategies towards faster quantitative imaging in atomic force microscopy. Eur. J. Control **11**(4), 384–395 (2005)
39. T. Osada, M. Godwin, International technology roadmap for semiconductors (1999)
40. S.Y. Chou, P.R. Krauss, P.J. Renstrom, Imprint of sub-25 nm vias and trenches in polymers. Appl. Phys. Lett. **67**(21), 3114–3116 (1995)
41. J. Strong, The Johns Hopkins university and diffraction gratings. JOSA **50**(12), 1148–1151 (1960)
42. D.J. Schroeder, An echelle spectrometer-spectrograph for astronomical use. Appl. Optics **6**(11), 1976–1980 (1967)
43. R.G. Tuli, Instrumentation in astronomy VIII (SPIE, Washington, USA, 1994), pp. 674–668
44. S.D. Berger, J.M. Gibson, New approach to projection-electron lithography with demonstrated 0.1 μm linewidth. Appl. Phys. Lett. **57**(2), 153–155 (1990)
45. H.C. Pfeiffer, W. Stickel, PREVAIL-an e-beam stepper with variable axis immersion lenses. Microelectronic Eng. **27**(1), 143–146 (1995)
46. A.N. Broers, J.M.E. Harper, W.W. Molzen, 250-Å linewidths with PMMA electron resist. Appl. Phys. Lett. **33**(5), 392–394 (1978)
47. P.B. Fischer, S.Y. Chou, 10 nm electron beam lithography and sub-50 nm overlay using a modified scanning electron microscope. Appl. Phys. Lett. **62**(23), 2989–2991 (1993)
48. K. Early, M.L. Schattenburg, H.I. Smith, Absence of resolution degradation in X-ray lithography for λ from 4.5 nm to 0.83 nm. Microelectronic Eng. **11**(1), 317–321 (1990)
49. N. Vorbringer-Doroshovets, F. Balzer, R. Fuessl, et al., 0.1-nanometer resolution positioning stage for sub-10 nm scanning probe lithography, in *SPIE Advanced Lithography. International Society for Optics and Photonics* (2013), pp. 868018-868018-10
50. G.R. Harrison, N. Sturgis, S.C. Baker, G.W. Stroke, Ruling of large diffraction gratings with interferometric control. J. Opt. Soc. Am. **47**(1), 15–22 (1957)
51. G.R. Harrison, S.W. Thompson, Large diffraction gratings ruled on a commercial measuring machine controlled interferometrically. JOSA **60**(5), 591–595 (1970)
52. G.R. Harrison, S.W. Thompson, H. Kazukonis, et al., 750-mm ruling engine producing large gratings and echelles. JOSA **62**(6), 751–756 (1972)
53. C. Mitchell, Diffraction grating fabrication in Australia. Spectrochimia Acta Part B, **54**, 2041–2204 (1999)
54. T. Kita, T. Harada, Ruling engine using a piezoelectric device for large and high-groove density gratings. Appl. Optics **31**(10), 1399–1406 (1992)
55. E. Kouno, P.A. McKeown, A fast response piezoelectric actuator for servo correction of systematic errors in precision machining. CIRP Ann. Manuf. Technol. **33**(1), 369–372 (1984)
56. J.F. Cuttino, A.C. Miller Jr., D.E. Schinstock, Performance optimization of a fast tool servo for single-point diamond turning machines. Mechatron. IEEE/ASME Trans. **4**(2), 169–179 (1999)
57. K.K. Tan, T.H. Lee, H.X. Zhou, Micro-positioning of linear-piezoelectric motors based on a learning nonlinear PID controller. Mechatron. IEEE/ASME Trans. **6**(4), 428–436 (2001)
58. M.F. Heertjes, X.G.P. Schuurbiers, H. Nijmeijer, Performance-improved design of N-PID controlled motion systems with applications to wafer stages. Ind. Electron. IEEE Trans. **56**(5), 1347–1355 (2009)
59. G.J. Maeda, K. Sato, H. Hashizume et al., Control of an XY nano-positioning table for a compact nano-machine tool. JSME Int. J., Ser. C **49**(1), 21–27 (2006)
60. H. Numasato, M. Tomizuka, Settling control and performance of a dual-actuator system for hard disk drives. Mechatron IEEE/ASME Trans. **8**(4), 431–438 (2003)
61. Leang K K, Devasia S. Hysteresis, creep, and vibration compensation for piezoactuators: Feedback and feedforward control, in *Proceedings of the Second IFAC Conference on Mechatronic Systems* (Berkeley, CA, 2002), pp. 9–11
62. J.M. Cruz-Hernández, V. Hayward, Phase control approach to hysteresis reduction. Control Syst. Technol., IEEE Trans. **9**(1), 17–26 (2001)

63. H. Shakir, W.J. Kim, Nanoscale path planning and motion control with maglev positioners. Mechatron. IEEE/ASME Trans. **11**(5), 625–633 (2006)
64. B.E. Helfrich, C. Lee, D.A. Bristow et al., Combined-feedback control and iterative learning control design with application to nanopositioning systems. Control Syst. Technol. IEEE Trans. **18**(2), 336–351 (2010)
65. S. Bashash, N. Jalili, Robust adaptive control of coupled parallel piezo-flexural nanopositioning stages. Mechatron. IEEE/ASME Trans. **14**(1), 11–20 (2009)
66. A. Sebastian, S.M. Salapaka, Design methodologies for robust nano-positioning. Control Syst. Technol. IEEE Trans. **13**(6), 868–876 (2005)
67. K. Kuhnen, H. Janocha, Inverse feedforward controller for complex hysteretic nonlinearities in smart-material systems. Control Intell. Syst. **29**(3), 74–83 (2001)
68. H. Shinno, H. Hashizume, H. Sato, Nanometer positioning of a linear motor-driven ultraprecision aerostatic table system with electrorheological fluid dampers. CIRP Ann. Manuf. Technol. **48**(1), 289–292 (1999)
69. Y.K. Yong, S.S. Aphale, S.O.R. Moheimani, Design, identification, and control of a flexure-based xy stage for fast nanoscale positioning. Nanotechnol. IEEE Trans. **8**(1), 46–54 (2009)
70. U. Staufer, in *Scanning Tunneling Microscopy II*. Surface modification with a scanning proximity probe microscope (Springer Berlin Heidelberg, 1992), pp. 273–302
71. C.F. Quate, Manipulation and modification of nanometer scale objects with the STM, in *Highlights in Condensed Matter Physics and Future Prospects* (Springer US, 1991), pp. 573–630
72. H.J. Mamin, R.P. Ried, B.D. Terris et al., High-density data storage based on the atomic force microscope. Proc. IEEE **87**(6), 1014–1027 (1999)
73. G. Binnig, H. Rohrer, C. Gerber et al., 7×7 reconstruction on Si (111) resolved in real space. Phys. Rev. Lett. **50**(2), 120 (1983)
74. B.D. Terris, S.A. Rishton, H.J. Mamin et al., Atomic force microscope-based data storage: track servo and wear study. Appl. Phys. A Mater. Sci. Process. **66**, S809–S813 (1998)
75. A. Davidson, MEMS-actuated magnetic probe-based storage, in *Magnetic Recording Conference, 2002. Digest of the Asia-Pacific. IEEE* (2002), pp. CE3–CE3
76. P. Vettiger, G. Cross, M. Despont et al., The "millipede"-nanotechnology entering data storage. Nanotechnol. IEEE Trans. **1**(1), 39–55 (2002)
77. E. Eleftheriou, T. Antonakopoulos, G.K. Binnig et al., Millipede-a MEMS-based scanning-probe data-storage system. Magn. IEEE Trans. **39**(2), 938–945 (2003)
78. M. Lutwyche, C. Andreoli, G. Binnig et al., 5× 5 2D AFM cantilever arrays a first step towards a Terabit storage device. Sensors Actuators A **73**(1), 89–94 (1999)
79. P. Vettiger, G. Cross, M. Despont et al., The "millipede"-nanotechnology entering data storage[J]. IEEE Trans. Nanotechnol. **1**(1), 39–55 (2002)
80. H. Pozidis, W. Haberle, D. Wiesmann et al., Demonstration of thermomechanical recording at 641 Gbit/in 2. Magn. IEEE Trans. **40**(4), 2531–2536 (2004)
81. M.A. Lantz, H.E. Rothuizen, U. Drechsler et al., A vibration resistant nanopositioner for mobile parallel-probe storage applications. Microelectromech. Syst. J. **16**(1), 130–139 (2007)
82. M. Lantz, H.E. Rothuizen, U. Drechsler et al., A vibration resistant nanopositioner for mobile parallel-probe storage applications[J]. J. Microelectromech. Syst. **16**(1), 130–139 (2007)
83. M.A. Lantz, G.K. Binnig, M. Despont et al., A micromechanical thermal displacement sensor with nanometre resolution. Nanotechnology **16**(8), 1089 (2005)
84. M. Lantz, H.E. Rothuizen, U. Drechsler et al., A vibration resistant nanopositioner for mobile parallel-probe storage applications [J]. Microelectromech. Syst. J. **16**(1), 130–139 (2007)
85. M.A. Rosa, S. Dimitrijev, H.B. Harrison, Enhanced electrostatic force generation capability of angled comb finger design used in electrostatic comb-drive actuators. Electron. Lett. **34**(18), 1787–1788 (1998)
86. J.B.C. Engelen, H.E. Rothuizen, U. Drechsler et al., A mass-balanced through-wafer electrostatic x/y-scanner for probe data storage. Microelectron. Eng. **86**(4), 1230–1233 (2009)
87. S. Devasia, E. Eleftheriou, S.O.R. Moheimani, A survey of control issues in nanopositioning[J]. IEEE Trans. Control Syst. Technol. **15**(5), 802–823 (2007)

88. A.J. Fleming, S.O.R. Moheimani, A grounded-load charge amplifier for reducing hysteresis in piezoelectric tube scanners. Rev. Sci. Instrum. **76**(7), 073707 (2005)
89. S. Parkin, X. Jiang, C. Kaiser et al., Magnetically engineered spintronic sensors and memory. Proc. IEEE **91**(5), 661–680 (2003)
90. J. Daughton, Spin-dependent sensors. Proc. IEEE **91**(5), 681–686 (2003)
91. Sahoo D R, Sebastian A, Haberle W, et al. Magnetoresistive sensor based scanning probe microscopy, in. *Nanotechnology, 2009. IEEE-NANO 2009. 9th IEEE Conference on. IEEE* (2009), pp. 862–865
92. S. Skogestad, I. Postlethwaite, *Multivariable feedback control: analysis and design* (Wiley, New York, 2007)
93. A. Pantazi, A. Sebastian, G. Cherubini et al., Control of MEMS-based scanning-probe data-storage devices. Control Syst. Technol. IEEE Trans. **15**(5), 824–841 (2007)
94. E. Eleftheriou, Nanopositioning for storage applications. Annu. Rev. Control **36**(2), 244–254 (2012)
95. A. Sebastian, A. Pantazi, S.O.R. Moheimani et al., Achieving subnanometer precision in a MEMS-based storage device during self-servo write process. Nanotechnol. IEEE Trans. **7**(5), 586–595 (2008)
96. C. Lee, S.M. Salapaka, Robust broadband nanopositioning: fundamental trade-offs, analysis, and design in a two-degree-of-freedom control framework. Nanotechnology **20**(3), 035501 (2009)
97. T. Tuma, A. Pantazi, J. Lygeros, et al., Tracking of high frequency piecewise affine signals using impulsive control, in *Proceedings of 5th IFAC symposium on mechatronic systems* (2010), pp. 90–95
98. T. Tuma, A. Pantazi, J. Lygeros, et al., Impulsive control for nanopositioning: stability and performance, in *Proceedings of the 14th international conference on Hybrid systems: computation and control. ACM*, 2011, pp. 173–180
99. T. Tuma, A. Sebastian, W. Häberle et al., Impulsive control for fast nanopositioning. Nanotechnology **22**(13), 135501 (2011)