Chapter 3

MACRO-SCALE PRECISION ALIGNMENT

3.1 Precision Machine Design Alignment Principles

Whenever two solid bodies are positioned with respect to each other, the quality of the alignment can be described in terms of the following two parameters: repeatability and accuracy. Repeatability is defined as the degree to which a part will vary its original position over time as it is being assembled and disassembled continuously. Accuracy is defined as the degree to which the part's position matches the desired position. Accuracy can only be achieved if the system's repeatability is good enough; however good repeatability does not ensure acceptable accuracy. Once the system's repeatability is acceptable, accuracy can be improved through adjustment and calibration.

Two basic principles, kinematic design and design for elastic averaging, are capable of providing high repeatability in the location of solid bodies to each other, beyond that obtainable by simple pins and slots [8]. These principles are used extensively by precision machine designers for the design of macro-scale systems. They may have an important application in wafer-to-wafer alignment of micro-scale systems, particularly in the MEMS and microelectronics fields.

3.1.1 Kinematic Couplings

Kinematic couplings are deterministically designed, static, structural couplings. In a deterministic system, the number of contact points between two solid models matches the number of degrees of freedom which are restricted. As the body is exactly constrained, its position can be determined in a closed form solution [8]. However, the point loads required by a deterministic system may cause significant Herzian contact stress on the couplings which limits its application. Repetability of 0.1µm has been reported with the use of heavily loaded steel ball and groove couplings. This material is subject to fretting corrosion, which requires wear-in and degrades the repeatability for high-cycle applications. Ceramic kinematic couplings are not subject to fretting corrosion and can be used with little or no wear-in [9]. The repeatability of a well-designed and preloaded ball and groove kinematic coupling is in the order of the surface finish of the grooves. Due to the low number of supports and high contact stresses, the stiffness of kinematic couplings is low compared to a surface-to-surface joint.

Kinematic couplings make use of concave features that fit into grooves. The shape of the grooves depend on the number of contact points that are required between the groove and the convex feature. If six degrees of freedom are constrained, one can choose to constrain three convex features in two degrees of freedom (DOF) each, as is the case with a three-groove kinematic coupling. Alternatively, one convex feature can be constrained in three DOF (i.e in a trihedral socket), a second convex feature in two DOF (using a regular V groove) and the last convex feature in one DOF (i.e a flat surface). Figures 3.1 and 3.2 show a three-groove kinematic coupling.

To ensure stability in a three-groove kinematic coupling, the grooves must be arranged in a triangular fashion, such that the normals to the planes created by the two contact points of each coupling, intersect within the coupling triangle [9] as shown in Figure 3.3.





Figure 3.1 Three-groove kinematic coupling disassembled

Figure 3.2 Three-groove kinematic coupling assembled



Figure 3.3 Coupling arrangement to ensure stability (figure by A. H. Slocum, *Design of three-groove kinematic couplings* [9])

3.1.2 Flexural Kinematic Couplings

Kinematic couplings can provide very high repetability. The price paid is a low joint stiffness, when compared to a surface-to-surface joint, and the fact that the surfaces of the joined parts don't mate. This is a drawback if the two parts to be joined are intended to seal. One way of increasing the joint stiffness and to allow the joined parts to mate while still achieving high repetability, is to mount the kinematic coupling elements, either the concave features or the v-grooves, on flexures, such that, when the coupling is lightly preloaded, it works as a regular kinematic coupling. As the preload is increased, the flexures bend until the surfaces of both bodies come into contact. The kinematic coupling is then fully pre-loaded and the rest of the load is taken by the mating surfaces. Although the repetability is slightly lower than the one in regular kinematic couplings, the joint's stiffness is increased significantly. This is the idea behind the "Kinflex" design, shown in Figure 3.4. [10].



Figure 3.4 Flexural kinematic coupling "Kinflex" (US patent 5,678,944 [10])

3.1.3 Elastic Averaging

Contrary to kinematic design, elastic averaging is based on significantly over-constraining the solid bodies with a large number of relatively compliant members. As the system is preloaded, the elastic properties of the material allow for the size and position error of each individual contact feature to be averaged out over the sum of contact features throughout the solid body. Although the repeatability and accuracy obtained through elastic averaging may not be as high as in deterministic systems, elastic averaging design allows for higher stiffness and lower local stress when compared to kinematic couplings.

In a well designed and preloaded elastic averaging coupling, the repeatability is approximately inversely proportional to the square root of the number of contact points [11].

Hirth or curvic couplings, used in serrated tooth circle dividers, shown in Figure 3.5, are examples of elastically averaged couplings. The serrated tooth circle divider uses two mating face gears. Both are the same diameter and have equal tooth geometry and tooth size. As the two face gears are engaged and preloaded, the teeth are lapped, the individual tooth size and position variations are averaged out over all the teeth, thus providing good repeatability [12].



Figure 3.5 Circle divider (figure by W. R. Moore, Foundations of mechanical accuracy [12])

Figure 3.6 shows a detailed view of the face gears disengaged. Figure 3.7 shows the same face gears engaged.

This type of coupling relies on stiff elements and requires large preloads. Furthermore these type of couplings often require a wear-in period to achieve very high repetability.

The principle of elastic averaging can also be applied to designs that use more compliant members, thus requiring a smaller preload. An example of an elastic averaged coupling based on low stiffness elements are Lego TM blocks.



Figure 3.6 Curvic coupling disengaged (figure by W. R. Moore, *Foundations of mechanical accuracy*[12])



Figure 3.7 Curvic coupling engaged (figure by W. R. Moore, *Foundations of mechanical accuracy*[12])

3.2 Elastic Averaging Bench Level Experiment

Elastic averaging can be used to accurately locate solid bodies, and may potentially play an important role in locating MEMS structures in a die or with respect to another MEMS device. To investigate this potential, a series of experiments were performed on LegoTM DuploTM blocks to qualitatively evaluate the repeatability that can be obtained through this principle. The press-fit assembly design of LegoTM blocks makes use of the elastic averaging principle, obtaining high repeatability [13,14].

Tests showed that the particular toy blocks used in the experiment, when assembled and preloaded effectively, have a repeatability of less than 5 μ m. It is anticipated that the actual repeatability can be improved from the one reported by better controlling the pre-load; nevertheless, the repeatability we measured is still quite impressive.

LegoTM blocks are prismatic, thin-walled, plastic toy blocks provided with projection or bosses symmetrically distributed on the top and bottom faces of the blocks [13]. Figures 3.8 and 3.9 show the top and bottom view of a 2x6 primary projection (PP's) building block. Primary and secondary projections are arranged such that, when the blocks are placed on top of each other, the primary projections of the bottom block engage with the secondary projections of the top block. Each projection engages in exactly three contact lines with its mating geometry [14], as shown in Figure 3.10. The dimension and location of the projections allows for the blocks to be press fitted on to each other [13]. The slight interference fit between the engaged projections of different blocks creates the necessary frictional engagement, or holding force, to keep both blocks fixed to each other [13].





Figure 3.8 Top view of 2x6 PP building block

Figure 3.9 Bottom view of a 2x6 PP building block



Figure 3.10 Cross-section at the interface of two blocks showing three line contact of every primary projection with adjacent secondary projections

3.2.1 Repeatability of a 2X4 Projection Lego[™] block

A series of experiments was performed on Lego[™] Duplo[™] blocks to determine the repeatability that can be obtained through elastic averaging on ABS injection molded parts.

The experiment consisted of repeated assembly and disassembly of A and B type blocks. Type A (96mm x 32mm x 19mm in size) and Type B (about 64mm x 32mm x 19mm in size). Type A block has 12 primary and 5 secondary projections. The shorter block (Type B) has 8 primary and 3 secondary projections. The position (sides and top face) of each block was recorded through every cycle, as shown in Figure 3.11.



Figure 3.11 Measurement target for repeatability experiment of 2x4 PP Lego [™] block

In a first set-up, the data was taken with a CMM. The same experiment was repeated using capacitive probes. Capacitive sensing was prefered because of its high resolution, repeatability, and accuracy (linearity) [15]. The resolution of the measurement system use in this bench level experiment is 5 μ m for the CMM and 0.05 μ m for the capacitive probes.

A gauge block, shown in Figures 3.12 and 3.13, was designed to mount the capacitive probes and constrain the bottom LegoTM block. The main requirements of the gauge block were high precision and low distortion. The design was chosen to provide a tight structural loop. Making the complete block one solid piece and directly probing the LegoTM block faces, minimized the Abbe error. The block consists of a central pocket to which the bottom building block has been epoxied. Capacitive probes are mounted on flexures on two faces orthogonal to each other.

Ejection pins were used to disassemble and assemble the blocks in order to avoid contacting the capacitive probe during assembly and disassembly. This set-up was needed because of the limited measuring range and the reduced clearance between the blocks and the capacitive probes. Four 3 mm bores were placed into the bottom block to give clearance to the ejection pins. Although the bores slightly reduce the bottom block's stiffness, it is assumed that it does not have a significant effect on the overall repeatability results.



Figure 3.12 Font view of gauge block used to measure the repetability of 2x4 LegoTM blocks

Figure 3.13 Detail of flexures to hold capacitive probes and ejection pins to disassemble the blocks

Capacitive sensing needs a conductive surface as a target, so a 25 μ m thick aluminum sheet was glued to each block as shown in Figure 3.14.



Figure 3.14 LegoTM block with aluminum sheet used as a target for the capacitive probes

The same experiment was repeated using chrome plated LegoTM blocks to eliminate the error introduced at the shim-block interface.

A routine was used to probe the block's position with the CMM in the first set up. In the second and third set-up, the output signal of the capacitive probes was connected to LabviewTM software through a data acquisition card and recorded for every assembly-disassembly cycle. The block's position was recorded once the readings had stabilized. Creep and thermal stress caused the readings to drift for abut two minutes. The output signals were normalized to the first read-out in order to eliminate any signal offset. The outlier measurements (maximum and minimum values) were dropped. The repetability was calculated as the range of the remaining data¹. The results of this experiment are presented in Table $3.1.^{2,3,4}$

Experiment	Bx [µm]	Tx [µm]	By [µm]	Τy [μ m]	Bz [μm]	Tz [µm]
CMM	5	19	5	20	5.3	20.3
Capacitive Using bonded sheet target	4.7	14.5	4.5	27.4	N/A	N/A
Capacitive Using chrome plated blocks	1.8	3.4	1.2	4.5	N/A	N/A

TABLE 3.1 Repeatability of 2x4 PP Lego[™] block

A "cap" test with the chrome plated blocks showed that noise in the measurement system accounts for sub-micron (10⁻⁷m) error. The cause of non-zero repetability of the bottom block is attributed to block deformation caused by the assembly and disassembly loads, as well as to thermal induced stress. This was confirmed by seeing a "growing trend" on the read-out of the probes over time, as seen on the plot in Figure 3.15. Some witness marks could be seen in the contact lines of the top block's secondary projections after the experiment had been repeated several dozen load-unload cycles. The data presented was taken from a short, 30 cycle experiment.

^{1.} Some authors define repetability as half the range. For the results presented herein, repetability is defined as the range of all data after eliminating outlier values

^{2.} Resolution of the CMM is $5\mu m$, resolution of the capacitive probes is $0.05\mu m$

^{3.} Repeatability results taken with CMM after 50 cycles; repeatability results taken with capacitive probes and bonded sheet target after 30 cycles, repeatability taken with capacitive probes on the chrome-plated blocks after 30 cycles

^{4.} Nomenclature after Figure 3.11



Figure 3.15 LegoTM block position in 30 cycle assembly-disassembly sequence, first bench level experiment

It was expected that the top blocks repetability in the y direction (Ty) would be better than in the x direction (Tx). The top block has 2.5 times more elements in the y direction than in the x direction, and repetability is inversely proportional to the square root of contact points. This however was not the case, and it is believed that since the assembly force was not carefully controlled during the experiment, the top block did not fully sit on the bottom block during some of the assembly cycles. The block's aspect ratio would cause a larger abbe error in the y direction than in the x direction, causing the unexpected results. In spite of this discrepancy, the repetability values obtained are quite impressive for these simple toy blocks.

3.2.2 Repeatability and number of contact points

A second bench level experiment was designed to evaluate the relationship between the number of contact points and the repeatability of an elastically averaged coupling.

The sequence described in section 3.2.1 was followed, but with a set-up that allowed the number of engaged primary and secondary projections to be varied. Six Lego[™] blocks, size 2x6 PP's, were epoxied between two Lego[™] plates, size 12x6 PP's, to create a relatively stiff monolithic block with 72 PP's, as shown in Figure 3.16. Two to five 2x6 PP's blocks were placed between two large monolithic blocks as shown in Figures 3.17 and 3.18. This modified the number of contact points between the blocks from 72 to 180.



Figure 3.16 Bottom and top view of the epoxied monolithic block used for the repetability vs. number of contact points bench level experiment

Three of the 2x6 PP Lego[™] blocks, which had previously been chrome plated, were used as targets for the capacitive probes. These target blocks were interconnected through a conductive shim embedded in the epoxied block.

One of the monolithic blocks was epoxied to a moving base, which in turn, was kinematically coupled to the base fixture via three canoe ball type couplings, as shown in Figure 3.19.

The base fixture consists of two main parts: a square block, which serves as a reference plane for X and Y measurements and a base with three press-fitted V-groove inserts, and a



Figure 3.17 Second bench level experiment using two 6x2 PP's blocks (72 contact points)



Figure 3.18 Second bench level experiment using five 6x2 PP's blocks (180 contact points)

pocket for a permanent magnet used to increase the kinematic couplings preload. The block constrains four capacitive probes using flexures.



Figure 3.19 Experimental setup for the second bench level experiment

A two piece, kinematically coupled fixture, as shown in Figure 3.19, is used to allow remote assembly and disassembly the monolithic blocks, without coming in contact with the capacitive probes. The capacitive probes are less than 1 mm away from the chrome

plated LegoTM blocks, and any physical contact with the probe while running the experiment causes drift in the read-out values. The top fixture can be tilted away from the capacitive probes to a safe distance for block assembly and disassembly. The kinematic coupling allows the moving plate to return to the original position relative to the fixture base with very high repetability. Canoe ball kinematic couplings have been shown to provide sub micron repeatability when subject to heavy pre-loads. The preload for the kinematic couplings in the bench level experiment is provided by the mass of the top fixture and two permanent magnets fixed to the top and bottom fixture. The repeatability of the kinematically coupled setup and the system's noise was determined through a cap test which consisted of repeated assembly and disassembly of the fixtures without disassembling the monolithic blocks. The cap test proved sub-micron repetability, the results of this test are presented in Table 3.2.

TABLE 3.2 "Cap" test for second bench level experiment

	Bx	By	Tx	Ту
Repetability [µm]	0.56	0.52	0.23	0.85

Thermal gradients as low as 0.5°C cause deformations in the aluminum fixture that exceed the repetability of the blocks. To avoid noise due to this source the whole system was placed in an insulating chamber and the position was recorded after the signal from the capacitive probes had stabilized.

The results of a 25 cycle run with 2,4, and 5 2x3 PP's blocks between the large monolithic blocks are presented in Table 3.3

As expected, both repeatability and standard deviation improve as the number of contact points is increased. Error theory predicts that the repeatability of an elastically averaged coupling is inversely proportional to the number of contact points. Although this is not reflected quantitatively, the experimental results clearly show this trend qualitatively.

Experiment	Χ [μm]	Υ [μm]	X Stand. dev	Y Stand. dev
2 blocks 72 contact points	8.15	10.95	2.484	2.759
4 blocks 144 contact points	5.47	6.23	1.271	1.737
5 blocks 180 contact points	2.805	3.59	0.768	1.021

TABLE 3.3 Repeatability results of second bench level experiment