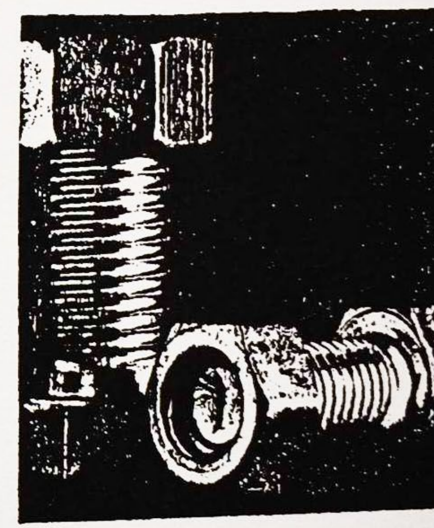
Notes

Precision Mechanics



David Kittell

Precision Mechanics

Introductory Dialogue 1

Rigid Body Kinematics 6

Flexure Kinematics 12

Rolling Kinematics 21

Adjustments 22

Holders, Clamps, and Locks 32

Structures 40

Configuration 44

Sensors 51

Preface

Over the last decade or so, the subject of mechanics as pertaining to the 'micro and nano', has been evolving as a distinct discipline. This is not small *parts* so much as small *displacements*. Instrument design has, of course, a great tradition and has been the hand maiden of many of the landmark scientific discoveries throughout history. Today, besides the obvious emphasis related to optical systems, the discipline has something to offer fields as diverse as microwave systems and atomic-scale microscopes.

This little book was inspired by a *Scientific American* 3-book collection of articles from its Amateur Telescope Making column. The spirit of this tradition was rarely mathematical or formal, chiefly for and by amateurs. An *amateur*, you will recall, is one who does something out of love. If he can get paid for it, so much the better! We will try to maintain that spirit here.

No particular claims for originality are being made. This is a collection of the most informal kind; it's neither complete nor well organized. Mostly, it's based on basement workshop experiences, simple, but hopefully, guided by a search for the underlying principle. If you're an old hand at this stuff, doubtless you will be at least entertained. If you're new to it, then welcome to an exciting, absorbing, and rather special world, with bright prospects for the future.

David Kittell
Stamford, CT 06906

March 1992

Precision Mechanics is the discipline of holding and positioning, in the world below a micron. We are concerned here with things like

- Stability over time and temperature
- Control of relative positions to an arbitrarily small dimension
- How to make neat mechanisms simply with low cost methods
- Intelligent use of (mostly) common materials and processes.

Our most valuable assets are

- Ability to sketch (anybody can learn to do it!)
- Close working relationship with a model shop
- Availability of a few basic metrology tools
- An open mind, and the courage to try new ideas.

Our units of measure, and their relationships are

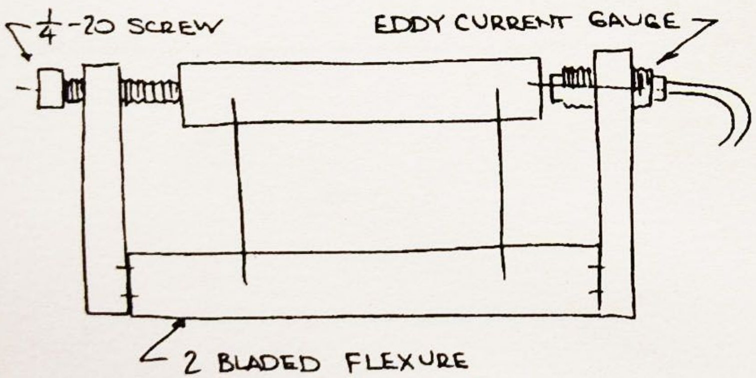
- 1 micron = 39.37 microinches
- 1 mil (milli-inch) = 25.4 micron (written µm)
- 1 micro-inch = 254. Angstrom Units
- 1 wave = 0.6328 microns (He:Ne).

Any self-respecting precision mechanic knows these units like the back of his hand, and can convert back and forth in his head with great ease and speed.

- This piece of paper is about 4 mils ≈ 100 µm.
- A high powered optical microscope runs out of steam at about 1 micron.
- We never write "µin" for micro-inch because it is too easily confused with µm.
- If you must use it, write it out. µm and µin differ by a factor of 40 to 1 !

If you are doing precision mechanics, your normal senses are practically useless. Thus our reliance on the dial indicator (about 1 µm resolution) and the eddy-current or capacitive gauge (about .01 µm resolution) to see what's happening.

The figure below is a cartoon of our two-bladed flexure demonstrator. This is among the simplest ways of providing a single degree-of-freedom (1 DOF) in translation. Those of you familiar with the categorization of such things may be surprised to learn that *DOFs are not necessarily straight line translation or pure rotation.*



As a warm-up exercise, let's explore our ability to adjust position using the ubiquitous 1/4-20 screw. Invoking our very best dial indicator, we find that with it's least division of 50 microinches (5/4 µm) we can't really see well enough to decide. (The dial indicator has insufficient 'gain' to resolve the capability of this crude screw! With the eddy-current gauge, we can now observe a limitation, probably about 1/2 µm.

Adjustment screws are such a common feature of many instruments that they are worth a close look. Some of the ways to improve on their precision:

- Use a finer pitch; this is okay up to around 80 tpi (threads per inch) but let's see what can be done with stock parts, wherever possible.
- Improve the bearing between screw and nut. Much can be gained here with only a little effort. We want a reasonably tight fit to minimize *sideplay*, which is a fickle source of instability.
- Put a long (say 1 to 2 inch) radius on the end. A squared off end will tend to bear on the edge. This can can have unpleasant effects.
- Use a much bigger knob if space permits. This encourages dexterity in the fingers and larger working leverage.

The two-bladed flexure is 1 DOF but it's motion is certainly not a straight line. More about this motion later. For now, it's a test bed for adjustment screws. Our 2nd experiment is to try the 4 suggestions for improving on the precision of the screw. To do this, we start with a standard 3/8-24 cap screw.

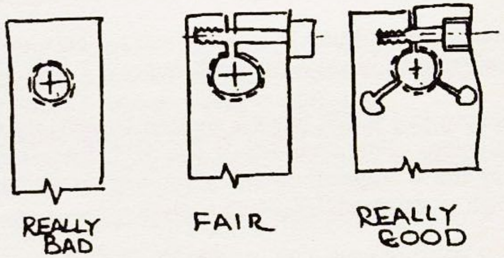
The radius is put on the end with a file in a lathe, then buffed on a hard felt wheel with emery. This gives a smooth bearing to the screw's output with minimal lateral influences. The threaded hole is definitely off the beaten track. We are exploiting the principle of the 3 jaw chuck or collet here.

The sketch below illustrates the collet approach. Note several features:

When the clamp screw is tightened, the adjustment screw is squeezed from three sides, and provides equal support in all directions. The classical method, labelled 'fair' is not nearly as enveloping, gives tight fit sideways but little tightening up and down.

The clamp screwhead is recessed. This tends to discourage meddling and gives a shorter screw with less tendency to cock as it is tightened.

The collet is made by drilling center hole, and smaller holes near the edges. A bandsaw is then used to create the slots. No accuracy is required here, roughly 120 degrees per jaw is ideal by not fussy. The small holes reduce stress concentrations and is usually a good way to end a slot. The tapped hole can be put in with jaws slightly 'sprung' with a screwdriver so that less clearance will be built in. We're trying to get rid of clearance!



If we put our adjustment screw into the adjustable collet and start to clamp down, we find that the screw has a considerable lobing action (turning resistance varies in a cyclic 'once-around' fashion). This can be cured quickly by lapping the pair together. A slurry of oil and 240 grit abrasive works well. The screw needs lapping far more than the tapped hole. One can lap many screws with the collet, keeping a firm clamp-down for rapid action. Make sure to remove abrasive before using to avoid any gritty feel and/or premature wearout. A small ultrasonic tank is handy for cleaning things like this.

Principle: If two parts bear on one another, and one of them is (nearly) perfect, then the pair will act (almost) perfectly.

However reluctantly, we must admit that one cannot rotate the screw without some clearance. This is generally true. If rubbing parts are involved, stick-slip is the inevitable result.

Also, we must put some kind of lube on the screw. Bare, especially clean, metals don't rub together smoothly for very long before destructive galling sets in. Unlike metals are better, e.g. brass against steel or brass against aluminum. Your author's favorite is molydisulfide grease, rubbed in with a toothbrush. A little goes a long way, and is amazingly effective for a long while, though not forever. More on tribology later.

A few digressions:

A screw is ideally 1 DOF. A helix can *only* do one thing, its position definable with a single number, say, 7.31 turns from a reference position.

The dial indicator is a mechanical machine with internal rubbing parts, and thus (jewels included) is sure to exhibit stick-slip. Common practice is to tap-tap-tap with a pencil to get a truer reading. In general, mini-shocks and/or low level vibration can be friendly, as they tend to reduce friction. You must anticipate this effect in designing and using precision instruments.

Principle: Stick/slip is a form of energy storage. Purposeful vibration is an essential method of reducing it in a controlled way.

The fine displacement meter and relatively coarse screw we have been using illustrates an important point in adjustments of all kinds. *The precision of the setting is ultimately limited more by the feedback device than by the adjustment machine.*

Stability of the adjustment, once made, is quite another matter. We will concern ourselves with these issues on a chronic basis. In fact, the success or failure of this discipline is largely determined by these two factors: adjustment precision and stability over time, temperature, and jostling.

When one part has to move w.r.t. another, something has to slide (pronounce 'rub'), roll, or bend. Our job is to select the best strategy for the application.

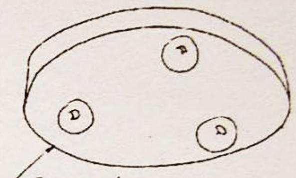
A few suggestions for becoming a better precision mechanic:

- Don't miss any opportunity to study how the other guy did it. Where possible, take things apart, feel for the fits, look at the materials.
- Learn how to draw *freehand*. It's your most valuable tool for visualization, discussion and selling of your ideas. Lines don't need to be straight, nor circles round.
- Use ink. Practice by drawing cubes and cylinders from all angles and in many sections. Artistic talent is not essential, but practice is!
- Keep a bound notebook of sketches of precision mechanisms, facts and ideas for future reference. It will become a highly valued possession. Draw small; you can get a lot more on a page.
- Experiment, experiment, experiment. Make breadboards to evaluate ideas before committing to sleepless nights. This will encourage trying novel ideas. Get your model shop guys to teach you to make models yourself. Hands-on experience remembers best. Analysis is fine, but there are usually surprises.

- Avoid the expensive and/or exotic. *Simplicity is at the heart of elegance and elegance is the first mark of an excellent design.* Search for it!

- Use bingo cards heavily. You should know every component available on the market, preferably first hand. This is a lifetime pursuit. Your standard component file takes routine maintenance but will be the envy of all your peers.

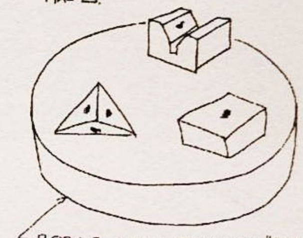
The term 'kinematic relationship' refers to a rather strict and formal concept. You may have trouble putting all the 'if/thens' to practical use, but it is an essential tool.



BODY A (UNDERSIDE) WITH BALL HALVES.

Simply put, any relative motion or position of one rigid body with respect to another can be described with no more than three translational and three rotational co-ordinates or 'directions.'

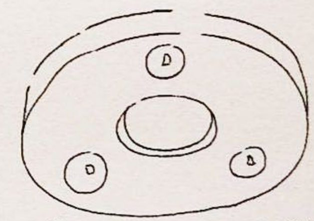
That being true, we can constrain body-tobody relationships by choosing constraints that limit relative movement. If these constraints (stops, hard points, etc) are chosen wisely, then the remaining freedoms will be as we specify.



BODY B WITH TRIHEDRAL "SOCKET", VEE GROOVE, AND FLAT.

The key point here is that *there should be no conflict to these constraints*, and thus no indefiniteness. This makes for stable, predictable, repeatable and thus precise relationships.

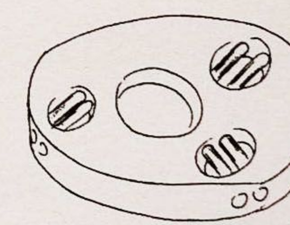
Of equal significance, and a fact often missed, the parts that do the constraining don't usually have to be very precise. A spherical surface doesn't need to be much more than smoothly convex, and a flat needs only to be not very curved but smooth.



SAME BODY A WITH SAME BALL HALVES

Illustrated above is perhaps the most famous of all kinematic arrangements. We describe it as follows:

1 ball in trihedral (3 sided socket)
1 ball in vee groove
1 ball on flat.



BODY B HAS 3 VEE BLOCKS

The lower picture is a more modern, and in some ways superior equivalent coupling. This one is described simply as: '3 balls in 3 vee grooves.' It's simpler to say and simpler to make.

In either of the variants shown, there will be 6 contact 'points' between the two bodies, no 3 collinear, and thus no degrees of freedom (DOF) remaining. To say that there are no degrees of freedom is equivalent to saying that they 'fit together' in a uniquely defined way, without any conflicting forces or allowable rocking, tilting, or sliding motions. The relationship is altogether comfortable, relaxed, and precise.

The surfaces that do make contact are best smooth with approximately the right geometry. One does not need any precision to create the parts. Only rather small contacts are actually involved. We refer to these contact areas as *bearings.*

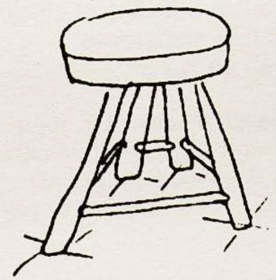
A *bearing* is simply where one part bears against another. Forget about ball bearings and the like for now. A bearing is a (usually static) contact spot between two bodies. If it's a kinematic relationship, the bearing points have been intentionally chosen to provide our choice of constraint vs. freedoms.

The pressure at these kinematic bearings can be amazingly high. You are advised to calculate the load and stress on bearing spots with the aid of 'Hertzian Stress Analysis.' Low modulus and high yield strength are of value here, as both factors help to avoid detenting. If contact stresses are too high, the material yields, leaving a pocket with a memory. This is no longer going to be kinematic, and all bets are off. Also, a little of Hertz will encourage use of largest possible radii for bearings. Roarke, or your Strength of Materials book, has a short but interesting section on Hertzian stresses well worth studying. Learn to use the general case of 2 bodies with curved surfaces. The ball in vee made up of two rods (cylinders) requires use of the general case to analyze.

The stool is about as good as any object known to illustrate rigid body kinematic relationships.

The 3 legger is kinematic because with its 3 bearings (constraints) it can do only the three things desired, and nothing else. I.e., on a reasonably smooth, even if quite curvy surface, it won't rock.

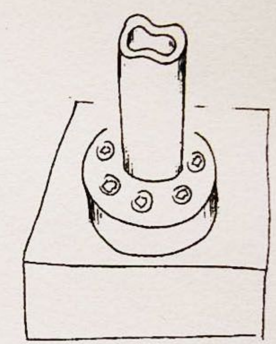
The 4 legger has trouble because it can't decide which set of 3 out of the 4 is going to do the defining. Since there are 4 possible sets of 3 legs as boss, the 4 legs invite indefiniteness.



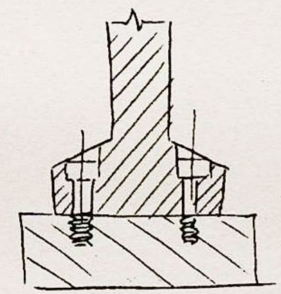
More generally, three bearings will form a plane which doesn't allow for non planar motions. One is then left with 2 translations and 1 inplane rotation remaining to control as required.

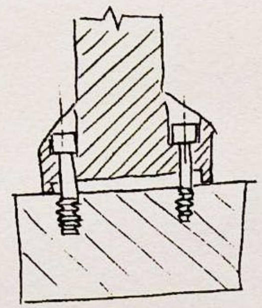
A rod in a vee groove is not kinematic. The reason is because the requisite 4 bearings are not really well defined for these shapes. The rod will touch the vee in 4 tiny spots, but who knows where? You must be able to identify the (kinematic) spots.

A slight divergence here while we're on the subject of bearings per se. Recall bearing is where two parts are in contact. We implicitly (and naively) think of a bolted joint as two like-shaped surfaces clamped tightly together, as at right.



The upper cross section reveals less than ideal conditions. Since surfaces are never perfectly flat, the bolts are going to pull the parts together, but an undesirable strain energy will be stored in the bent metal. Bending is not as benign as straight compression/tension, because it can raise havoc as it releases, especially with temperature cycles. Also, if not otherwise relieved, most of the bearing will be in the immediate zone of the screws, little elsewhere.



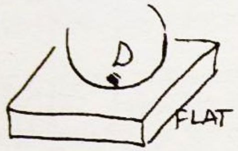


The lower cross-section shows an undercut on the flange. This practically insures that the bearing will be at the widest possible radius. Slightly out of perfect bearing is more gracefully forgiven with this technique. It's more trouble, but the resulting bolted joints will be stiffer, more stable, and more precise.

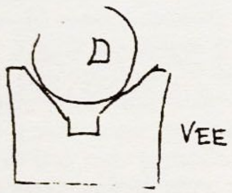
If you insist on the best possible bolted joints, then relieve just around the bolt holes and lap the parts together. This is old fashioned but still hard to top. Lapping two parts together is easy and fast, and quick to test. Ask your model maker. Lapped together surfaces also have the lowest thermal resistance. This is important where temperature gradients are a factor in your stability budget.

While you have your wrench out, tighten the screws! Most people have no idea really as to the right torque for a screw. More about this later when we get to fasteners.

The closely associated topic of structural resonances and damping enter here. Common engineering materials have far less inherent vibration damping then is generally believed. In anything less than a weldment or monolithic assembly, damping comes essentially all from the joints and mounting interfaces. More on this under resonances and damping.



1 CONSTRAINT,
5 DOF



2 CONSTRAINTS,
4 DOF

3 CONSTRAINTS,
3 DOF

CONE SUBSTITUTE
FOR TRIHEDRAL

These are the basic elements we were discussing in kinematic bearings. The same parts used in many other combinations show up often in rigid body mechanics. The ball-on-flat requires that the tip of the ball comes into play. If this is the lower most part, it will get scraped or abused and will take a little extra protection. The flat is probably best quite hard, and can be a glued on pad. Crazy Glue is good for this because of the zero thickness bond line, although I find it somewhat fickle. More and more of the world is being glued together, seemingly without undo compromise.

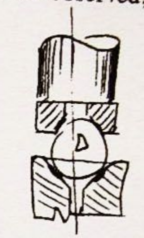
The ball-in-vee is used in many places. There are several ways to make a vee. Dowel pins are of precise diameter for press fitting into a simple reamed hole. Also they're hard and low cost. Experiment with fits. Use lube.

No one has pronounced that the included angle of the vee must be 90 degrees. This is the angle of choice only where the designer cares equally about vertical versus horizontal stiffness. A more open angle is stiffer vertically and softer horizontally, etc.

If the vee is made in some other way, look at the cross section where the ball actually touches to assess the included angle. If there are reasons for the ball to actually slide in the vee with any regularity, especially with high contact stresses, a flexure or rolling element should be considered. Static (or nearly so) pressure points are the most difficult bearings to lubricate. Again, cyclic thermal or high vibratory forces can scour a contact surface enough to cause a loose part or assembly.

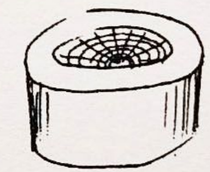
The purist loves the ball-in-trihedral. This approaches the ideal 'ball joint' so highly coveted by us all. It's not too handy a pocket to make however. A thru hole in the bottom makes cleanliness easier. Bearings should be clean, yes, but a small amount of some dry lube is highly recommended for lower friction and more durability of the surfaces.

Under most circumstances a less pure but easier to make poor man's trihedral is the ball-in-cone. This works beautifully if certain precautions are observed, as follows.



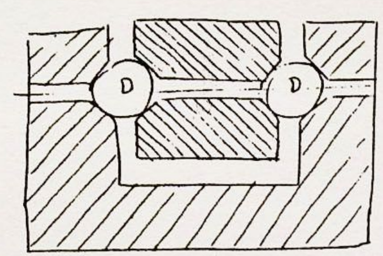
SQUEEZE
BALL BETWEEN
CONES TO INDENT
SPHERICAL SEAT.
ACTUALLY, A
NARROW ANULUS.

A nice cone is not made with a center drill or countersink. Better to turn on a lathe with a tool bit. If the ball-to-cone annulus is going to see big stresses, or much relative motion, it may help to broaden the contact line into a spherical area. Use a spare ball and squeeze in a milling machine vise or arbor press. Don't overdo this else you will cause more than localized distortion which may be disastrous. A narrow annulus is best. Again, a little lube is essential if the joint is going to do much moving under load. Choice of materials is worth some experiment. I often use a tungsten carbide ball in a brass or soft steel cone, but lube can be a problem.



WHERE POSSIBLE,
SPECIFY SINGLE
POINT TURNED
CONE.

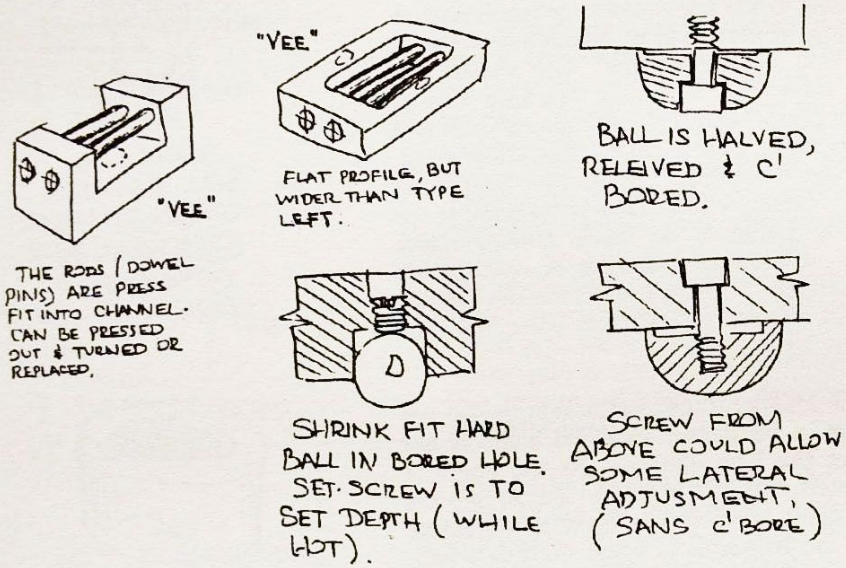
In like manner to our discussion on the vee groove, a 90 degree included angle is not always the best choice. As the angle is narrowed, the ball begins to wedge into the opening. This can be used to advantage when the ball-cone is being used as a ball-joint and rotation about the cone's axis in not desired, e.g. a tapped ball in a cone is usually superior to the old reliable spherical washer sets. A 60 degree angle resists torque about the thread axis. This acts like a self-aligning nut.



The 'perfect' hinge is an illustration of several points. A grade 10 ball (10 μinch sphericity and mirror-like surface) costs peanuts, but is close enough to perfect that when combined with some elastic averaging it is capable of defining a precise axis of rotation. To do this, we put two ball-and-cone pairs in series. This isn't quite kinematic because (think trihedral) there are 3 constraints on each ball for 6 total bearing points, while allowing for axial rotation.

The conflict can be largely removed by making one, or possibly both of the outside legs with axial spring compliance. This can be made an adjustable spring force for 'tuning.' What are we tuning for? It has either a residual play, or there's too much friction to resist for high finesse positioning. Alas, this is the fate of all bearings that depend on a rubbing contact. Coulomb friction is our major foe.

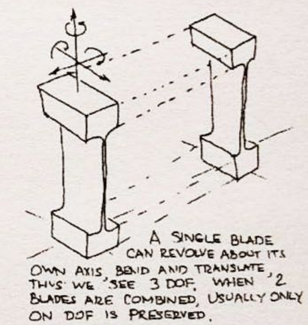
A few practical alternative vees and balls are sketched below. Your Hertzian analysis will prompt you to choose a hard ball. The shrink fit approach has been used successfully but is bulky. A press fit is the next choice for a hard ball. And we understand that one can anneal a hardened ball, machine, and then reharden. We tend to use cold worked 302 cres. as a compromise.



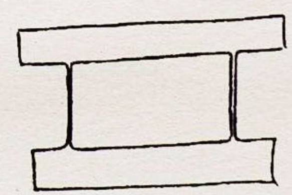
All of the contact joints discussed so far involve rubbing friction. To null out or otherwise deal with this friction may be a task. An experiment with a few pound load on 1" diameter ball/vees showed a repeat

For small displacements and angles, there are sometimes good reasons for using bending action rather than rubbing. Bending substitutes a spring rate for the friction of rubbing. For precise control, bending is preferred because spring forces are linear, dependable, and often quite small around zero deflection. Rubbing friction is highly non-linear and maximum under static conditions. This makes it impossible to move arbitrarily small distances. It takes considerable force to move any distance if rubbing is involved. When the force is high enough to actually achieve movement, a sudden breaking-loose is experienced combined with a jerky stop. Try adjusting your rear-view mirror a tiny amount. You will find yourself in a limit-cycle around the perfect aim point. Stick/slip results from Coulomb friction, especially when static is higher than dynamic friction, the usual case. A flexure-based relationship is (ideally) completely free of this stick/slip hysteresis effect. An incremental change in force gives a linear increment in displacement; we know this as Hooke's law. From a control viewpoint, this is highly favored.

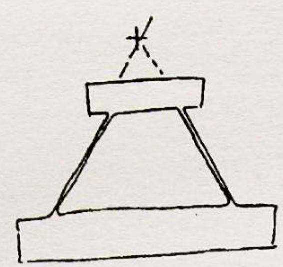
We will use our kinematic relationships to determine DOFs much as we did earlier. The guide to understanding a flexure-based coupling is to examine each flexure element for constraints and DOF's. For example, the single-blade flexure has 3 constraints that obviously result from its reluctance to bend within it's own plane. If we take 2 blades with their ends tied rigidly together, they will constrain each other's DOFs, except in the one quasi-linear direction. The motion is parallel but *not* a straight line. This *is* a single degree of freedom though.



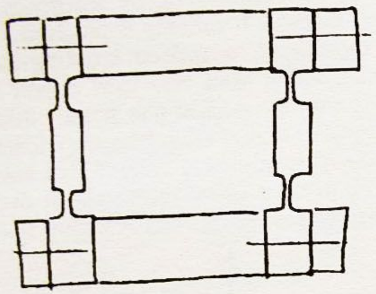
The other trick to understanding flexure systems is to determine their *virtual* intersection(s). The two *parallel* blades *cross at infinity*, and thus constrain the (non-grounded) body to rotate about this point at infinity, which is translation (to a 1st order at least).



If we apply this rule to a *non-parallel* pair of blades that (virtually) intersect at a *finite* point as is the case at lower left, then they will create an axis of rotation about that point. For angles of rotation of 1 or 2 degrees, this makes a neat remote pivot.

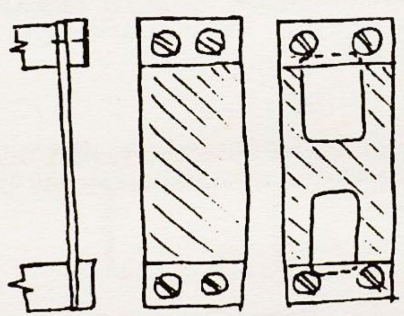


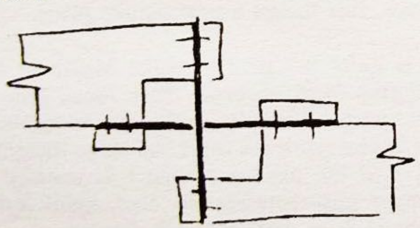
It is quite practical to make blades in the integral form shown. This saves lots of little parts. There is a little trick to machining blades from the solid like those illustrated, and the choice of mat'l is usually a compro mise between low cost, lightweight body parts vs. high strength flexures.



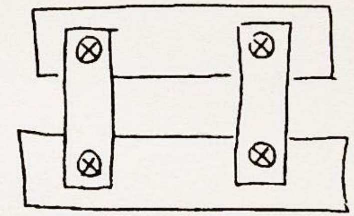
Often we see a 'blade' machined out of the solid. This is much stiffer in the stiff directions and less prone to column buckling. We're less apt to have ideal material properties in the thicker initial section. The carved blade can be secured to the frame with less joint loss. One can also combine carved blades with integral construction, i.e. the whole mechanism is machined or investment cast as a *single* piece. Occasionally one may need the lowest possible joint loss. A particular case is where the flexure is providing the spring for a resonant device like a scanner, where you want both the moving mass and the restoring spring all made out of a single piece. A weldment is second best, but avoid putting the (electron beam) weld near a high flex zone so as not to reduce the spring's metallurgy or fatigue strength.

Also popular is the uniformly thin blade but with windows. This will reduce the spring rate while preserving stiffness where it counts most. A stress analyst will help you decide the best shaped windows. Use radii in the corners if long fatigue life is important. One ends up wanting a high strength but low modulus material that is easy to fabricate. Beryllium copper (CA 172) machined in the 1/2 hard rolled form, then heated in air for 3 hours at 600 degrees F. makes good blades. The details of attachment have a large effect on the resulting spring rate because the blade is apt to pried open its clamp and/or slightly yield them in the corner of the joint. Proper bearing and hefty screws that are fully torqued-up help here. Pins in addition to the screws may also be wise if high shear forces are anticipated.





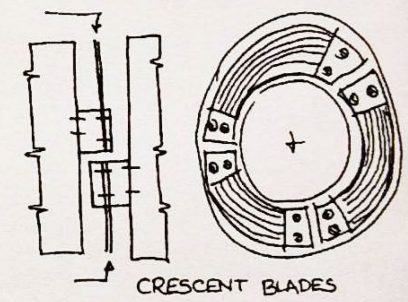
This is also a classic (cross-bladed) flexure arrangement. Often, 3 blades are used with a common *virtual* intersection to balance things, taking the form of two half-wide blades on either side of a full-wide blade. The precise axis of rotation *does* shift with bending angle in this design. The heart of the ubiquitous 'Bendix' pivot is constructed with crossed blades very much of this sort.

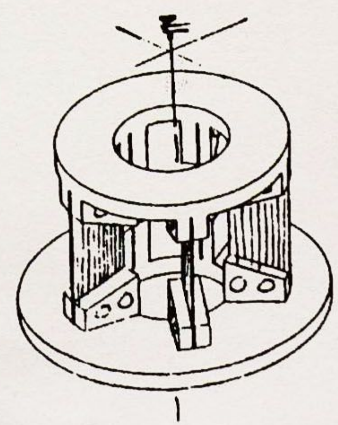


This one is a four-bar linkage made with Bendix pivots. You recognize it as being kinematically equivalent to our previous two (parallel) bladed flexure. Four bars have great versatility; they are way under utilized. They can take one kind of motion and transform it into another, especially over small distances and angles. These flex pivots are more expensive than blades if you're going to talk real quantities.

The crescent flexure is so called for fairly obvious reasons. With 3 blades arranged around a circle in this fashion, alternate ends being tied to alternate bodies, the device constrains all planar motion, leaving 3 DOF, two tilts and Z. With 2 of these separated by a tube and grounded at the outside ends, we have a Z only DOF, which is without peer for a focusing mount.

This compound crescent flexure does have a subtle rotation about the Z axis when it is displaced along Z, but optics are rarely sensitive to this. If carefully made, one can focus a 0.4 n.a. lens over +/- a millimeter with less than arc-second tilts. The focus mounts in CD players are a near cousin of this design and can be made with compactness and low spring rate. The helix direction must be phased on the ends, to allow them to rotate (however subtly) in the same direction.

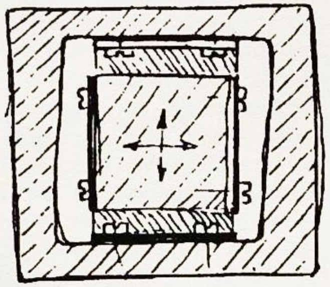




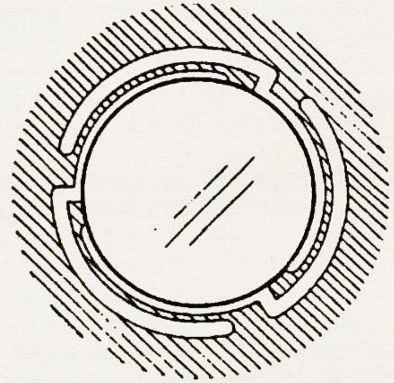
Called the paddle-wheel flexure, this has only 1 DOF. We can guess this immediately by observing that the center of symmetry is in common with all blades. This can be made in a large range of sizes. It can be used alone as a rotary table (+/- a few degrees) or in pairs to create a long axle. A hefty limited angle gimbal suggests itself. The paddle-wheel bearing suffers shortening along the axis, as you would expect, though it rarely matters.

Any of these multiple bladed flexures are not inherently conflict free because high locked in stresses can be created when they are assembled.

Called a 'box flexure,' this is just a pair of two-bladed flexures mounted one on top of the other, at right angles. One produces an upward turning sag curve, while the other curves downward. Ideally, parallel-bladed arrangements are driven along their neutral axis, which is normal to the blade direction half way between top and bottom. This avoids unbalanced moments that tend to cause subtle tilts. The picture shows a top view of the box flexure. One set of blades goes downward to a plate, while a second set comes up the other sides. Thus they are in series or, as we say, they are *stacked*, even though one set is nested inside the other.



This ring mount is comprised of three radially flexible blades that will support an optic in a kinematic way. Your author makes these out of one piece by first boring a hole in a solid disc, then relieving the area around the hole with a small endmill on a rotary table. The optic is then epoxied by injection through 3 radial holes drilled through from the outer diameter to the pad surface. The pads can be relieved with a die grinder to form a glue pockets. Don't let the glue form a meniscus, as this can cause local distortion of the glass. The blades can be bent by hand to increase or decrease the spring clamping force as desired.



Analytically, this is a cousin of the 3 ball/3 vee-groove kinematic mount. Each blade is free to move in a radial direction, same as each ball can move in its radial vee.

In this one, we are once again making a vee groove; a cone stacked on a 2 bladed flexure. This provides the same kinematic action but without any rubbing or stick/slip. You can equip the center of the cone with an air line to rest the ball effectively on a film of air pressure. One could also pull a vacuum on the same line to create a clamping action between ball and cone. This has been used in practice with great success. The cone seat must be accurate, smooth and clean for reliable vacuum. Removal and replacement with 3 of these between 2 bodies can be ultra precise.

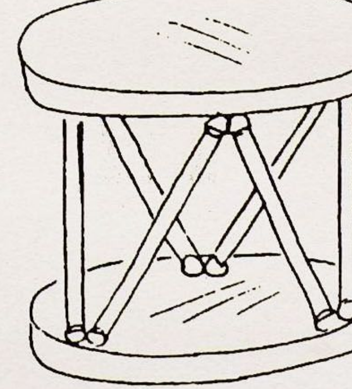
We call this 2 and 1/2 DOF because it doesn't really twist very well (has high torsional spring rate). But yes, it does act like a ball-joint since it can rotate about all three axes. These are almost always used at both ends of a rod or tube to form a single constraint between bodies. This allows freedom in the other five directions. Shops don't like to make necked-down blades like this smaller than 15 or 20 mils thick. Also, you will often need a high-strength material as the MC/I rises quickly with angle.

2½ DOF BALL JOINT

It also suffers in compact arrangements because angular flexibility takes blade length. This in turn moves the blades away from the end, where they are less effective. Also, the centers of rotation are in different places for the two bending directions.

The lower structure is somewhat of a maverick to categorize. There is one ball-jointed leg for each of six constraints and no remaining DOF's. The legs can be made of tubing with ends beveled to meet the balls at 45 degrees. If holes are drilled thru the balls, extension springs can be used to load the joints.

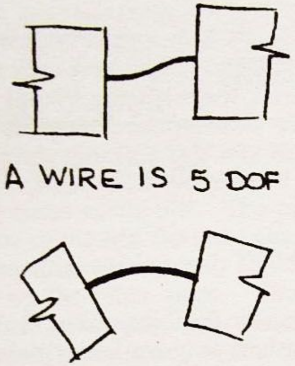
This is a simple, low cost, light-weight and symmetric structure. Electrical conduit tubing of low carbon steel makes for low expansion. Not too obviously, it can be made into a 6 DOF system by installing a variable-length drive in each leg. By suitable change in length of the legs, any combination of motions can be had. Try one of these with tennis balls and cardboard mailing tubes for fun. The tubes can telescope so that the individual legs can be adjusted to demonstrate their effects. In practice, each leg does the same thing, but left or right handed.

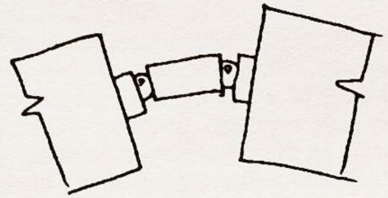


THE STUART (STEWART ?) TRUSS
6 DOF

To make it do any arbitrary move, takes a little matrix inversion, which is a function of both where you start and where you want to go. The plate end of the balls can be sunk into a cone for simplicity.

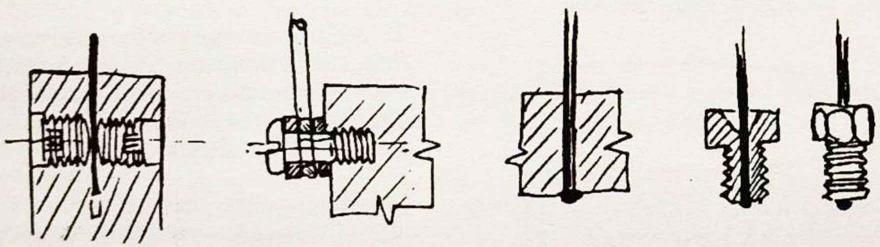
The famous Link pilot trainer uses a structure of this type with hydraulic pistons for legs. By controlling the relative lengths and rates of change, one can create an exciting ride. It is a natural to think of precision mechanics within a narrow context of instrument design or the like, but many of the underlying ideas have broad applications, from micromachined parts to skyscraper and bridge design. There is in fact, a company that specializes in flexures for bridge construction. A feel for 'bearing' can be important whenever parts are to mate in a controlled stress, stable way.





A DOUBLE BALL JOINT - 5 DOF

Another basic flexure element is the wire. It's often desired to make a connection between A and B in such a way that only one freedom is constrained. For many purposes, a length of wire has essentially the same kinematic properties as a double ended ball-joint. The wire is preferred in a precision situation because only linear, noise free forces and torques are generated with relative motion of A and B. Any two-piece ball-joint of conventional design presents an impossible trade-off between lost motion and high coulomb friction-based stick/slip. Incidentally, adding plastic liners does reduce the friction but plastic can extrude under pressure showing creep, instability and thermal problems.



For wire flexures we often use hard drawn, torsionally straightened 301. The yield strength is pushing 200,000 psi, and it's readily available. Of the many techniques seen to terminate a wire, these illustrations show our favorites. Of critical importance here is to avoid annealing the wire with heating. A hex-head cap screw of the same material is counterbored (to increase the free length of wire), drilled through, and reamed. The best fit for the wire is as tight as can be lightly pressed into place. Before e-beam welding, bake the assemblies, to destroy hydrocarbons. Use short weld cycle to minimize the total heat input.

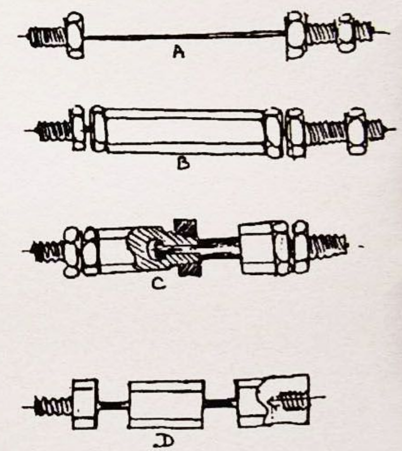
Still referring to the last picture, we have low-tech and high-tech methods of terminating a wire flexure. Whatever the case, a tight fit for the wire in the hole is desirable. Forming a loop and screwing down seems crude but works well for light-load situations and is fast & cheap. Clamping between two long radius-nosed set screws also works well in larger sizes. E-beam welding does the least wire damage and has been satisfactory in practice.

Almost always ball-joints show up in pairs; one on each end of a bar. This is necessary for 5 DOF. A single, short wire is less successful as a 3 DOF because it lacks lateral stiffness. More about 3 rotations later. Wire flexures made with welded-on screws are handy and broadly useful. By spacing the screw-heads close together, a bending limit is built into the flexure. One must think through the installation sequence to avoid over-stressing the wir

Of the types shown, the long, straight wire has the lowest bending and torsional rates but is limited in column buckling. By replacing the single wire with 2 short lengths at each end, we have a better axial stiffness situation. Hex stock, or even a tube (for light weight) will serve to connect the ends.

Type C uses an in-line collet for length adjustment. The collet is made by splitting the hex body 3 ways, terminating each in a generous hole. The tapered clamp-threads for the collet are made with a hardware-store pipe tap and die.

The collet version also solves the blind hole problem faced by the A and B types. Notice A and B have a nut shown on one end. The left end can be screwed into a blind tapped hole but the right end must then pass through an open hole to access a nut. The type C link is easiest to install because of its adjustability. The collet can be set to slip on an overload that sometimes occurs during assembly.

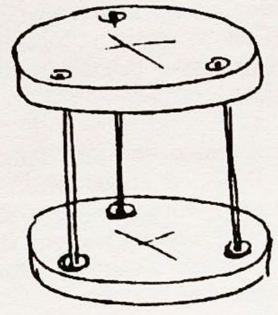


A SMALL FAMILY OF 5 DOF CONNECTING LINKS.

Hex heads and connecting lengths are practical as they provide wrenching flats for installation and adjustment. We have used these of wire diameters from .036 to .070 inch and could be scaled much further in both directions.

The type D link is machined from the solid. This is straight-forward to make and use but lacks any limit protection and must be heat treated material, such as a precipitation-hardening alloy (Be:Cu or the 17-X PH steels perhaps). This trades off simplicity for adjustability. Make some tensile and compressive load tests on breadboard flexures to verify your analyses for peace of mind!

The three parallel wirelink table allows the top to move around in the two dimensional world. The links can be made by any of the ways shown on the last page. This has the same X and Y translation as the box flexure with axial rotation thrown in. Aligning a target requires these three adjustment freedoms.

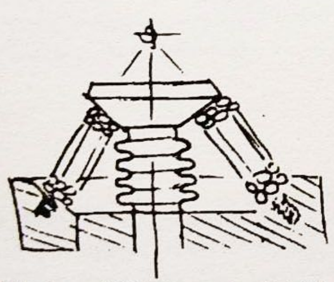


AN X-Y-θ PLATFORM

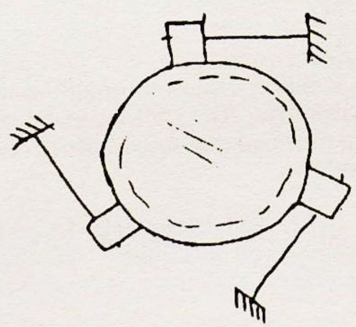
This can be made with the lower plate brought up to the top level with a tube, putting input and outputs into the same plane. We still must contend with the foreshortening (z axis) displacement that comes with moving the top plate in any of the three coordinates. The analytic assumption of a ball-jointed link is accurate for small displacements.

If we proceed with *non-parallel* links, then we will still have 3 DOF but of modified form. For example, the three links positioned on a conical surface will produce a 'ball-joint of ball-joints!' This makes a practical 3 rotations about a remote point design. If *axial* rotation isn't wanted, a central bellows will resist torsion. The bellows is a 5 DOF and good for broad precision mechanics situations.

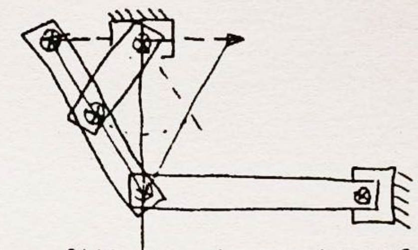
By attaching the links in tangential directions, we form the compliment of the parallel link platform. Here, the 3 DOFs are into the paper, axial translation and tilts about the in-plane directions. This one is appropriate for supporting many things in the planar direction. Growth and shrinkage of the central body is kinematically accomodated with this design and is, once again, analogous to the 3 ball/3 vee for in-plane motions.



AXIAL ROTATION CAN BE STOPPED BY USING A CENTRAL BELLOWS



3 TANGENT LINKS



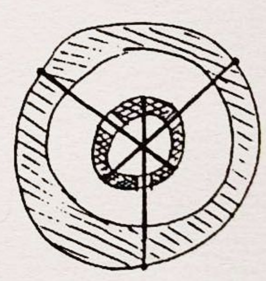
ANOTHER 4 BAR LINKAGE WITH 1 DOF TRANSLATION BUT OF MORE EXTENDED TRAVEL CAPABILITY.

The linkage pictured above doesn't really belong here, but it has worthwhile properties. In order to make a useful translator, at least 2 of these would be tied together. In general, we must fall back on the rules for success when trying to combine various devices with constraints and DOFs. One must use care to see that there is sufficient constraint but no contradiction, with no more than one limiting any given DOF. With 2 linkages in parallel like this one, the potential for conflict must be somehow resolved.

The concentric rings with connecting wires shown below is particularly interesting because, in addition to being a 1 DOF pivot, it can demonstrate the property of spring-rate compensation. With the inner and outer members being made of rings, a *tensile preload* may be imparted to the wires. Because of the balance of forces, the natural bending forces of the wires in rotation can be nicely balanced by these preload forces, with their over-center action. No spring rate means no torque required to turn.

There are many similar schemes for compensating flexure spring-rates. In fact, one is always dealing here with a stability concern, because the compensation can potentially cause a *negative* net spring-rate, in which case, the device would toggle between two symmetrically displaced angles about null.

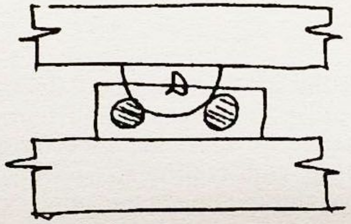
Also, careful analysis and observation will show that the precise change in force (or torque) with rotation has a differing shaped transfer function for the wire's bending versus the change in radial overturning force opposing it. This can produce rather startling results. Also, temperature extremes would be likely to cause trouble with the balance. This device as shown is apt to be used *in pairs* on either end of an assembly where limited axial rotation is needed. The compensated spring-rate could be adjusted to establish a torsional oscillation. An added part may be required to better determine the axial position of the center suspended parts.



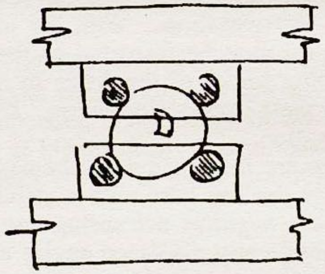
A SMALL ANGLE 1 DOF ROTATION

Much of the world's machinery rolls on balls. The antifriction bearing, ball or roller, operates in a different realm than precision mechanics. The old saw about 'nothing rolls like a ball' may be true, but ball and roller bearings have problems of their own. Recall that we faulted the ball/vee bearing because of stick/slip due to rubbing. True enough, static friction in a ball can be three orders of magnitude lower since it can roll, but we now must contend with a loose part. By adding a *second* vee and using a rolling ball instead of a single vee and rubbing ball, the friction is very much lower. The joint is still kinematic, but alas, it is no longer precise! Precision, you will recall, insists on high repeatability. We can no longer control the precise position of the ball or its orientation within the vees. A kinematic joint made with a rolling element is doomed to uncertainties of position, of a complex, difficult-to-control nature.

The second aspect of rolling-element bearings to appreciate is their overconstraint. A rotating or linear ball bearing has 10's or even 100's of balls competing for action. This is good for stiffness and/or durability, but cannot qualify for kinematic. Both ball and roller bearings for linear and rotary motion have this overconstraint implicit in their nature. True also for ballscrews and ball-splines, if the races are preloaded heavily enough to remove all play, then these bearings will run with uneven force or torque requirements and run a noisy imprecise path (at the submicron level.) Remember, we are interested in behavior in the sub-micron world. Utter repeatability and smoothness are paramount to our goals.

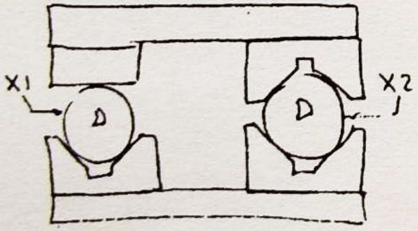


RUBBING VEE



ROLLING VEE

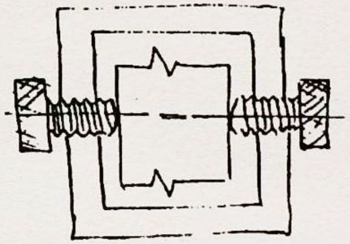
The schematic below is of a linear slide that is kinematic, although not entirely repeatable because of the loose balls. Okay, you need to travel a greater distance than can reasonably be accomplished with flexure-based systems and you need low friction. This method of a kinematic design based on rolling elements can be used successfully.

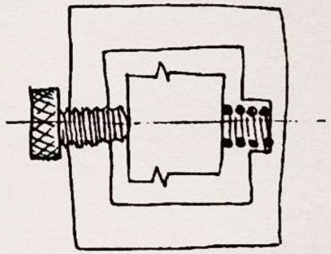


The lack of precision lies in the imperfect nature of the parts. And since we can lap the vees and balls to sub-micron accuracy of form, it is possible to prevail. The saving fact here is that the action on the parts is not overconstrained and is capable of frictional coefficients of .01 or better. Localized, short term repeatability, which often is most important can be very high. A retainer or guide may be needed to keep things where they belong.

Hopefully, our earlier experiment with the 1/4-20 screw persuaded you of the value of good feedback when making any adjustment. We provide an adjustment to put something in the right place. It may be a hand adjustment, or it may be a servo-driven mechanism, but in either case, the feedback signal is crucial. Is it noisy? Is it available in real time, on a continuous basis, without perceptible delay? Does the feedback signal have accuracy, resolution or precision problems of its own?

All of the criteria we use for qualifying precision mechanical devices is equally applicable to feedback devices, with additional look-outs such as stuff like parallax error and Abbe error. Is the signal presented to the operator completely honest, or is he being misinformed? At any rate, one is unlikely to make any adjustment more accurately than the feedback device warrants.



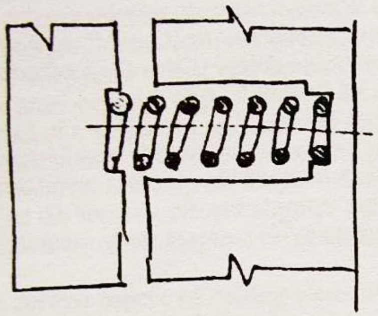


Often, particularly with flexure-based schemes, the adjustments may have interactions. All is not lost however; we will develop a sequence that takes us from most interactive to least as the adjustments are made. It's not unusual for a sequence of adjustments to be iterative, for the above reason.

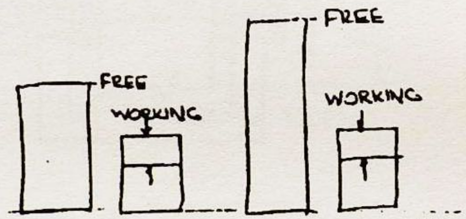
Then, having made a suitable adjustment, we are concerned with its stability with time, temperature and disturbances. The left figure exemplifies rigid adjustments, e.g. lathe chucks. The adjustments are overconstrained and thus compete for dominance. Its practical value lies in the resulting stiffness of the setting, provided the screws are tightened the right amount. This is poor machine design, where the operator is asked to decide. *Your* success depends on *his* skill. In the hands of a qualified operator, the N-jaw chuck can be exquisite.

Using an opposing spring rather than a second screw is much the most common practice. The screw should have at least 1, and preferably 2, diameters of bearing in the tapped hole. The nose is best rounded slightly, even polished. The mating surface should be somewhat harder than the screw if the adjustment will see appreciable use, to avoid indenting, tracking or memory effects. To some extent, the adjustment will be easier with a bigger knob and a finer thread. Factors of 2 improvement are hard to achieve for practical reasons. Does the setting change when tapped? If so, you have some work to do. Submicron adjusts are easier to make than to maintain because of sneaky stick/slip energy storage.

Having decided to use an opposing return spring for an adjustment, we want it to act perfectly elastic. This means avoiding contact with the spring except at the ends. A one-coil-deep pocket at both ends is ideal. The hole can be slightly undersize to capture the spring with no scraping on the edge of the hole. The longest practical spring will provide the most uniform force over the range of adjustment because the fractional change in length is less. The longer spring has a lower spring rate, but is compressed to the same working force. Column stability puts an upper limit on length, sometimes before space limitations and should be checked.



What is the right choice of spring force? Too much overloads the screw, making the adjustment task more difficult, and increases rate of wearout for thread and nose bearing. Too little force gives a soft feel and stability problems due to insufficient dominance over the mechanism involved. Don't hesitate to try several wire sizes on both sides of your best guess. A precise feel is psychologically important to the user's confidence in your machine and must be earned with extra effort on your part.



LONGER FREE LENGTH MEANS MORE UNIFORM WORKING FORCE

Principle: It's usually the things we haven't thought about that cause trouble!

A great deal can be learned by studying the mechanisms and devices of others. The discipline of Precision Mechanics has a centuries-old tradition that sees only slow change. It's rare for a new device or material to come available with a strong impact on the field. A precision mechanism will be largely ignored if well done, otherwise it, and its designer, will receive substantial uninvited attention. Quick fixes are rarely as satisfactory as an adequate initial design. Wise engineers make models and breadboards sooner and oftener.

There is a rule in manufacturing that says "The fewer adjustments, the better." This stems from 3 factors:

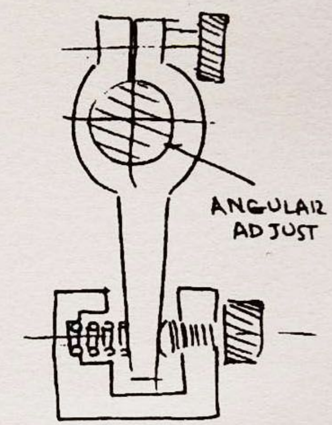
· Adjustments cost money to tweak.

· Adjustments can be abused by non-qualified tweaters.

· Adjustments are not known for great stability.

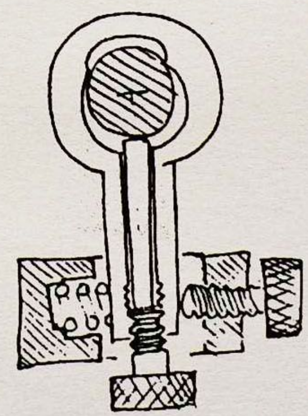
A classic in our field, the tangent screw mechanism deserves some study. We want to turn a shaft in precise increments. Often the range of the tangent screw and/or arm are insufficient and therefore, require a locking device to disengage the fine motion. With the clamp unlocked, the shaft is free to turn to about the right place. The clamp is then engaged, and a fine adjustment is performed.

It may be obvious that the precision of the adjustment will be lost if the shaft is not properly supported. Lost motion or high stick/slip in the shaft bearing will defeat the linearity and stability of screw adjustment.



If the adjustment is implicitly also being used as the subject shaft's lock, then perhaps a stronger return spring is appropriate.

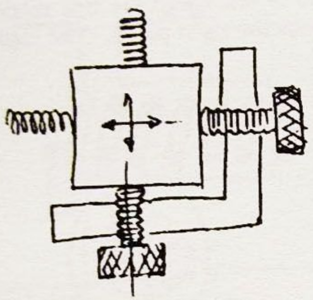
The tangent screw works on the end of a lever arm. As such, it acts like a big knob with a very fine pitch. With an arm of a few inches and a fine pitch screw, a gain of 1000 is not unreasonable. The 'gain' being the turns ratio of knob to shaft. This brings up the question of range versus resolution in a control.



If the adjustment range is large, lots of turning may be necessary with a long arm and a fine pitched screw. This can be tedious. Also, a tangent working over a range of +/- 5 or at most 10 degrees is a practical range without problems due to oblique line of action. Note the +/- sign. Ideally the screw acts at right angles to the lever arm. This is always a good idea; to have a force acting normal to a lever. Also, don't ignore the side force associated with the unbalanced moment created by the lever. It is not a couple.

The left version of the tangent screw uses a different locking device. This is handy when the shaft end of the arm isn't easily accessible. The clamp screw better have a softer nose that the shaft to avoid detenting over time. Putting the threads on the knob end is easier to construct.

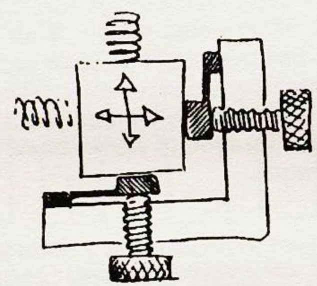
Recall the box flexure. We want to provide an adjustment in each of the two DOFs. A major look-out here is the interactions between the X and Y screws.



BAD PRACTICE

Notice that we will cause a rocking of the X screw when we adjust Y. Lateral forces on the screw-ends raise havoc on cross-axis coupling. Any lateral bias in the screws will be relieved with a shock load, and the setting will be spoiled. Also, the lateral shear on the opposite screw imparts a torque to the payload. In a situation where rotation is stringently to be avoided, lateral forces will have to be eliminated.

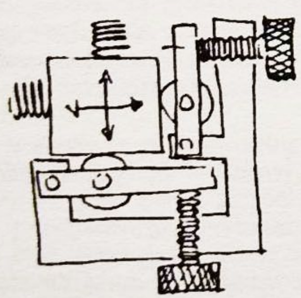
The second version of the box flexure drive solves the problem of side loads on the screw end by interposing an intermediate leaf-mounted follower that isolates the screw and table top. This is reasonably effective when the counter spring forces can be low. These springs have to overcome the box flexure's spring rate, and are proportional to the desired range of motion. Thus by keeping both of these at a minimum, the torque due to 'scraping by' will be minimized. Use of the lowest possible friction material between rubbing parts can help.



SOMEWHAT BETTER

Even better is to reduce the side forces by introducing a 'live' follower. The wheel and axle are not perfectly round, nor is the pivot without lost motion. This scheme can work with substantially higher bias springs.

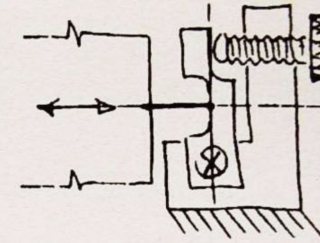
Significant here is the use of the versatile lever. The lever does several good things for us. Note that it isolates the drive screw from the load. Lateral forces don't couple easily. The lever is providing a 'gain' of 2:1 as sketched. The effective pitch of the screw is doubled, but alas, the range is cut in half. This contradiction comes with the lever.



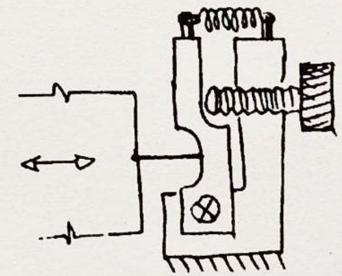
EVEN BETTER

Where the cross-axis travel is small, we can replace the uncertainties of the wheel and axle with a wire link follower. Notice the lever has been contorted to put the adjust screw bearing, the wire end, and the Bendix pivot all on the same centerline. The relaOtionship among pivot, wire and screw can be interchanged to give more or less lever reduction, or a different exit point.

The lever bearing wants to be wide based to resist side loads coupled in through the screw. One could also counter-bore both lever and payload to increase the length of the wire. If it is a 2 or 3 DOF situation, then the length of wire directly affects the cross coupling between adjustments. Note that there are no loose parts here which is apt to provide higher precision and stability. We are still depending on an external spring (not shown) to load the mechanism.



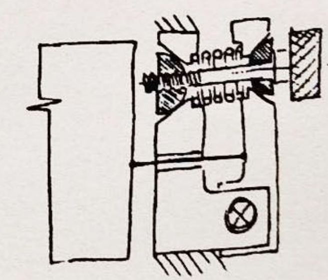
WIRE FOLLOWER



LOCAL RESTORING SPRING

Better to use a localized spring if practical. In this way, the forces associated with the positioner are operating over short distances with less distortion to the payload, which is probably already burdened with flexure spring forces. This positioner is self contained. It's somewhat better to keep the restoring spring at a lever distance where the displacements are high so that a smaller spring will do its job. Other things being equal, the design with lower forces will be the most stable. Creep is stress induced; lower stress means less creep. Attention to detail pays off when building an actuator for micro-positioning. This design is capable of 0.1 μm precision if produced and used with care.

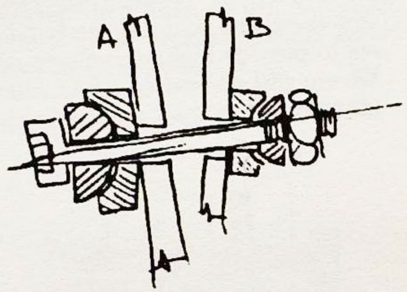
Finally, we would prefer a compression spring really centralized at the screw as shown. The thread is now put into a threaded ball-half and cone. This gives the screw much better lateral support over a longer wheelbase, and puts the lever on the outside for a longer connecting link. The tapped ball-half should go into a 60 degree included-angle cone. This provides a wedging action for antirotation. The two threaded parts and thrust washer should be made of differing materials and preferably lubricated in some way for smoothness and long life.



LONGER WIRE LINK
PLUS LOCAL PUSH SPRING

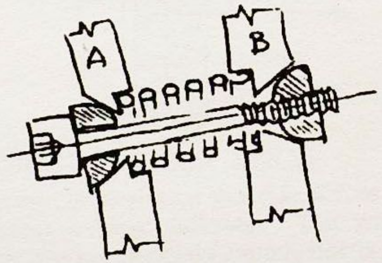
The philosophy for this adjuster derives from the basic need to control the spacing between two objects. A crude start in the right direction begins with a screw and a pair of spherical washer sets. A and B are constrained only along the direction of the screw. The washer sets allow for some misalignment and attempt to provide reasonable bearing. With much of an angle between A and B, there is a tendency for the screw and washer sets to slide down the wedge. Spherical balls don't fit spherical seats in a kinematic way. The contact zones will move around, and in general don't really locate the axis of rotation. We throw out the concave washers and replace with cones formed integral with the A and B parts. Now, a good quality ball will fit the cones nicely, and define an axis that is stable.

We can mill away at least half of the ball at each end. One could argue about the head of the screw bearing on the ball half as being slightly indefinite radially with a finite clearance. If the ball is fixed to the screw then the ball must rotate in the cone with attendant concerns for the local tribology.



SPHERICAL WASHERS

By making the back cone a 60 degree included angle, the back ball can be made to serve as a self-aligning nut. There is a subtlety here; if A rotates with respect to B about the axis of the adjustment, then the screw will tighten or loosen! In fact this has never shown itself to be a real problem. We have also included a counter spring to the set for several practical reasons. An adjustment wants to be stiff enough to dominate its local situation. By making the spring coaxial, motions or adjustments in the other DOFs won't cause any change in the local force. Try putting the spring anywhere else in a multi DOF arrangement; you will quickly see why. Mentioned earlier, the spring must be chosen to have sufficient force to handle the load extremes when the adjustment is at a maximum length. Use a longer, softer spring with more working compression for steadier force.

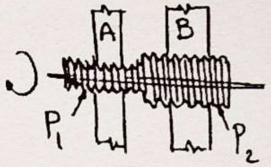


A BETTER WAY

Making the cone in one side and a shallow counterbore for piloting the spring in the other does take thicker pieces of material at least locally, and is more trouble, but well worth it. Compare this picture with the one on the last page. Note here, we are using just the bare bones of the prior adjuster. This stripped-down version is more appropriate for initial set up, or factory alignments done once or rarely. A spot of epoxy on the back end will hold the adjustment for posterity. Later, we will add a differential feature to this basic device for higher resolution. Incidentally, while setting a gap or making an adjustment, you will insure a more stable result by tapping it down.

One of the major causes of instability in a screw adjustment is lateral wobble. We have been showing the head of a cap screw passing thru a hole in the adjustment end. The screw can be fixed to the ball. This is good because play in the clearance hole will be eliminated and side wobble will be reduced. It is bad because the ball half must now turn in its cone, which it can do, but at a somewhat higher torque penalty. Supporting a relatively long screw at its ends with the push spring coaxial has low lateral play, and must undergo less angular accommodation as the ends experience differential motion due to the other potential 5 DOFs not controlled by the particular adjustment here.

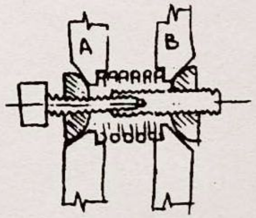
When we attempt to increase the adjustment's resolution, we can use the lever, and/or we can use a finer thread. But how fine is reasonable? Industry sees 80 tpi (threads per inch) even 100, but these are tiny threads, easily buggered, and not terribly accurate.



A DIFFERENTIAL THREAD ACTS LIKE A VERY FINE PITCHED SCREW.

Another alternative for high resolution is the differential thread. A good deal can be accomplished without leaving the domain of the stock room. With no special parts, we can get about as fine as you are likely to need.

The screw has a different pitch on each end. If it is rotated 1 turn, it will thread into A by a distance $1/P_1$. It will also thread into B by a distance $1/P_2$. The change in separation between A and B is then the difference (differential) between them.



$$\frac{1}{P_{eq}} = \frac{1}{P_1} - \frac{1}{P_2} \quad \text{or,} \quad P_{eq} = \frac{P_1 \text{ X } P_2}{P_1 - P_2}$$

HERE'S A DIFFERENTIAL VERSION OF OUR STANDARD ACTUATOR. NOTE BALL HALVES HAVE A TAPPED THRU HOLE FOR LOW LATERAL PLAY.

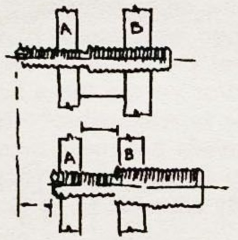
We also speak of the differential screw's 'gain', which is the increase in resolution over that of P_2 acting by itself.

For example, with the 1st screw a 10-32 and the second a 1/4-28, the equivalent pitch, Peq is 224 tpi, for a gain (over the 10-32 by itself) of 7.

The screw is made from 2 long setscrews or from threaded rod. The 10-32 threads into the 1/4-28 nicely. When we apply this to our ball/cone screw, it now acts as if it had a screw with 224 threads per inch. This corresponds to a lead of 113.4 µm/turn.

Having gained all this extra resolution, we must be prepared to pay for it in two ways. Notice how we can drive the screw over its entire range, but the net displacement of A with respect to B is quite small. The gain reduces the available output range by the same factor.

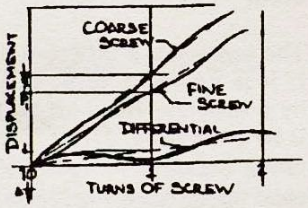
Any differential thread needs to be loaded in one direction to be useful. The reason being the otherwise embarrassing amounts of backlash. The two threads are in series as far as play is concerned. It can easily take several turns to reverse a load without sufficient bias.



THE GAIN OF THE DIFFERENTIAL ALSO REDUCES THE OUTPUT RANGE BY THE SAME FACTOR!

Using ordinary rolled fastener threads, a gain of 5 to 10 is often the best trade-off between errors and resolution. The graph shows the way pitch errors add up. You are going to see a random phase between the two thread run-outs.

It is not uncommon to experience an actual reversal of direction in the net displacement for a pair of threads with too close a ratio. Not only will it take endless turning to get anywhere but the trip will be a drunken stagger! Note that in the graph, where we have subtracted the motions, there is actually regions of the difference with a negative slope. This is going *backward*.

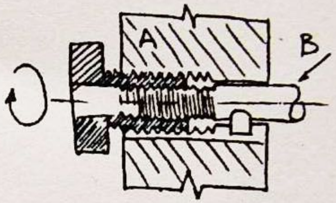


UNFORTUNATELY, THREADS HAVE ERRORS. WHEN WE ARE RIDING ON THEIR DIFFERENCE, THE ERRORS ARE STILL ONCE AROUND BUT CAN BE ENORMOUS.

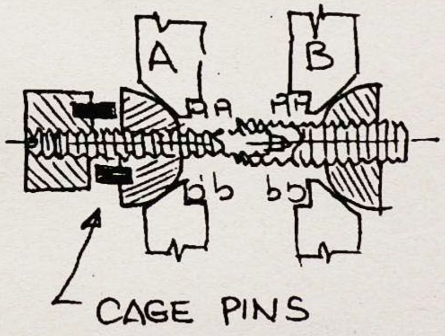
Another version (of many) is handy for a precise length adjustment. If the two threads chosen were 3/8-24 for the outer thread, and metric ISO M6-1.0, then we would have a gain of over 17, for a differential pitch of 435.4 threads per inch. This is high resolution (58.3 μm/turn) but the range is only 1/17th of what we had. This is less than the random difference between their thread starts making it impractical to put together with a controlled overall length.

Using typical threaded parts, the resulting lead errors in the differential are easily perceptible, although not reversing. In a differential arrangement, it is necessary to provide an anti-rotation device between two of the three parts. See figure next page. This will also add to the effective backlash. In fact, even with a one way spring bias, some softness may show up on reversals due to a finite clearance between the pin and keyway. A bellows could both load the screw and provide lash free antirotation over the limited range available. In the differential version of our standard actuator, we are depending on something else in the completed assembly to provide the antirotation between A and B. Since we are likely to use this for controlling a 1 DOF lever, the rotation is taken up by the lever's pivot so all is well with no extra keying requirement.

As a practical example of too much 'gain', with the combination of pitches M6-1 and 3/8-24 (or 5/16-24) we have a gain of 17.86 and a equivalent pitch of 453.+ tpi. If the full travel of the screw is 1/2 inch, then the differential has a total range of only .028" which is not even 1 threads' worth of random assembly, thus we can't ever count on being able to set standard length. One way of getting around this if you really don't need range but do have to set an initial position is to build the mechanism pictured to the right inside a tube that can be slid and clamped to zero. Don't forget the preload spring between A and B.



ANOTHER VERSION OF THE DIFFERENTIAL. NOTE THE PIN & KEYWAY.



Still not satisfied with having to swap too much range for extra resolution, we add one more wrinkle to the completed adjustment mechanism.

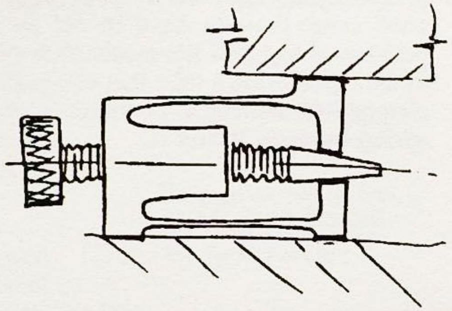
We add a pin to both the knob and the upper ball half for a cage. Now the differential action takes place with +/- one-half turn. Beyond that, the pins engage for a direct drive of the screw into the rear ball half. This is just a direct action screw for a rapid traverse. When the knob is then reversed, it acts differentially again. In practice, the automatic switch between coarse and fine is hardly noticeable.

Note, the knob is held onto the end of the threaded rod with a jam set-screw. This locks on the knob in a neat coaxial fashion.

When considering an adjustment, a knowledge and appreciation of the intended user is invaluable. Whoever it is, you can be sure that he doesn't care about the beauty of your design. He wants to make his setting and have it stay set indefinitely. Don't expect users to read instruction manuals or detailed explanations. The ideal design is *intuitive* for the user, and he will hardly notice it because it works so well.

Occasionally one will see a fine adjustment with rather heavy grease. Having some lubrication is essential, that's good. But the velvety feel of a highly damped mechanism caused by the grease can lead to unacceptable drag effects. Drag has a time constant that can be annoyingly long.

This gadget can be made in any size and is good for adjusting a spacing in close quarters. The screw has a taper turned on one end of 15 or 20 degrees included angle. The body is milled with integral flexure top and bottom, with the taper acting between sides of a slot cut in the end. One screw can be put into the bottom at the knob end to fasten it down, although it's most effective as a loose part. Also, a ball can be pressed into the top and bottom output lugs for some uses. Backlash will be experienced in the screw because there isn't anything loading it. High friction tends to make it reasonably stable.



The topic of adjustments is a vast one. But there really are a few underlying principles at work. In your author's experience, there are often surprises, not all pleasant, when trying novel ideas. The breadboard is always appropriate to check out a brilliant concept.

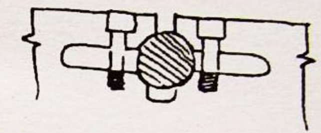
It's tempting to use elastic compliance as an ally in this context. The most common example of this is probably the practice of selective tightening of screws that compete with one another for dominance in bending. This is a sure sign of overconstraint. Poor stability is the likely outcome.

The best way to learn is to build breadboards and measure the result. This is itself somewhat of an art since time and money are always scarce. Just make sure your tests are stringent enough to uncover the limitations of your design in the context of its intended use. Resolution, backlash, windup and stick/slip are all easy to quantify with a micro-displacement gauge. Stability is quite a different challenge. Drift over both time and temperature need consideration. Some things to look for:

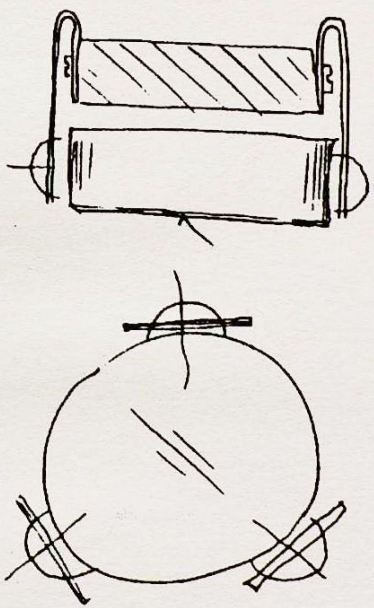
> Does it follow the same path going up in temperature as it does on the way back down? A cardboard or styrofoam box and lightbulb heater can work well, especially hooked up to a Variac. A few temporary T/C's judiciously placed will help take the guess work out of time constants.
>
> Drift can be elusive to track down because of implicit assumptions made in identifying potential culprits. Quite often the test set-up has more drift than the device under test.
>
> Don't forget to include vibration and shock in your suite of tests. It's a rare adjustment that won't respond occasionally, sometimes disastrously.

Pay attention to any "tuning" that may have to be done to your brainstorm in order to achieve full performance. Usually the designer isn't around when production tools up for building assemblies in quantity, and they don't want to fuss with 'filing to fit,' or selective assembly.

If you use rods or tubes for structure, they often appear along the edge of a plate or bar. Here is, again, an integral joint where the plate is milled out (or cast in) to form a chuck. This works best when the round parts are straight. A set of plates coupled this way can get overconstrained badly. The plates may be made as a matched set. A reamed hole of the best-fit diameter should be put into place before the milling. We have used these clamps on fused silica rods to make a long, low expansion instrument at low cost. Rods of this kind are cheap. They're sufficiently straight for structural purposes, but brittle.



CLAMPING A ROD OR TUBE CLOSE TO AN EDGE.



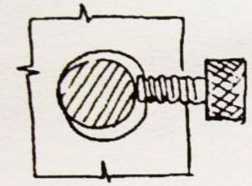
SMALL PARTS LIKE MIRRORS CAN BE HELD SEMI KINEMATICALLY.

This plate holder is, once again, a relative of the 3 ball/vee groove 0 DOF mount. The spring clips can be bent around by hand to secure the best compromise between squeezing the plate vs. residual play. The attractiveness of this mount is its compactness and ease of interchanging multiple plates. The ball halves want to be either radiused or undercut. A 1/2 " dia. ball for a 2" dia. mirror works quite well. Keep the excess glue to an absolute minimum to avoid meniscus effects. Use the mount itself for a glue fixture.

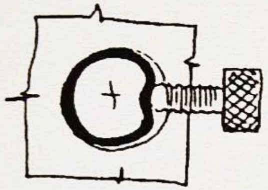
The back plate is apt to have a large thru hole, if for no other reason, for light-weighting. The ball should fit about 1/4 of a diameter into the blade. Use a countersink on the hole so that there is no sharp edge contact. This technique has been used in an interferometer with good results. Leave an extra length of blade sticking out beyond the front for a handy finger lift to remove and insert mirrors. Always better to avoid the need for tools when possible.

In considering size, what works nicely at a 2" diameter, may be disappointing at 4". Twice diameter means 4 times surface area and 8 times volume and mass, for a proportionate change. This is a large change in scale!

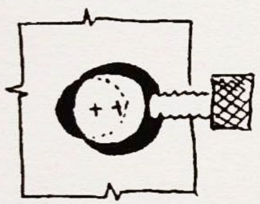
Often a source of mystery and disappointment is the shaft-in-hole problem. It seems obvious that the hole must be bigger than the rod, else they wouldn't go together. When a screw is brought up against one side, it is likely to be rigid in the plane of the rod and screw. In the other direction, there can be fatal rocking. This can cause trouble.



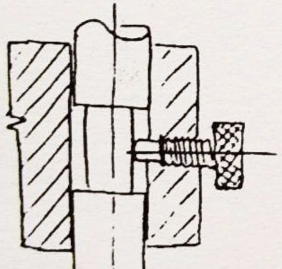
An important exception to this rule is in clamping a thin-walled tube, such as a Bendix flexure pivot. It goes without saying that a precisely reamed hole will put the tube into a reasonably tight (read push) fit, in which case the (round nosed) screw will cause deformation of the tube. This has an all-over tight locking action. Too tight on the screw and it's curtains.



For most applications, a double hole is the best. To make this, an undersized hole is drilled with an offset of about 1/4 diameter then the desired size hole is bored with an endmill. The hole size matters nothing now since the rod only contacts the holes at their intersections. Think of it as a poor man's vee-groove. With a little extra touch these edges can be relieved over the central half of their length. The joint will then be kinematic. The screw can have a soft nose, even with an integral spring. This offers a hint to the operator to avoid over tightening.

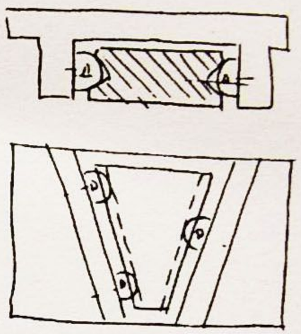


A knob is preferred to a slot or socket if there is danger of overtightening. An ordinary set screw has a cruel end and will bugger an unhardened shaft. The vee block hole can have lots of clearance, so that a bummed shaft will still go in and out, not being as likely to hang up.



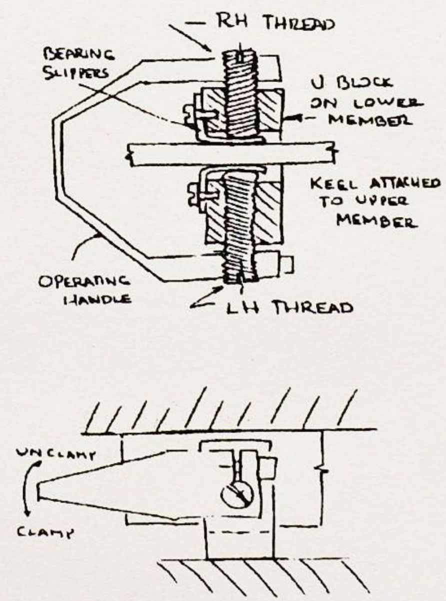
Consider an undercut shaft for situations where you want to intentionally limit the range of motion. The shaft is much less likely to become a loose (say lost) part. The sketch exaggerates the depth of undercut required. One still needs to avoid marking the shaft with the screw because marks will cause unpleasant memory effects.

Not too many years ago, can openers were held to the wall with a keystone-shaped plate on the removable opener, and a like-shaped socket that stayed on the wall. This can be made kinematic.



KEYSTONE SEAT MAY BE KINEMATIC

Note that each ball half contacts the slot at two points, for a total of 6 constraints. This suggests a compact interchange mount with high potential precision. Potential, because only attention to detail will give a fully satisfying result. The geometry is best longer than wide. An included angle of 20 to 30 degrees is okay. Avoid large overhanging loads, such as can openers. If used for a sample holder, a magnet at the bottom might help insure adequate and repeatable seating force.



A CLAMPING MECHANISM TO DEMONSTRATE A FEW IDEAS

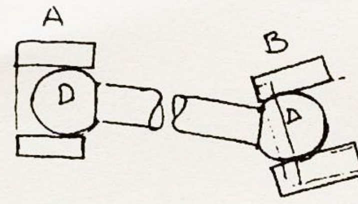
Clamps and locks often seem to cause more problems than they solve. Can you possibly avoid using one? On a 0.1 μm adjustment, it will take a rather sophisticated lock that won't disturb the setting. The one pictured here attempts to avoid some of the common pitfalls.

The moving member is extended with a 'keel' that fits within a slotted block in the fixed member. Left and right handed screws are tightened in against the keel when the lever is pressed. Notice, the clamping action is at right angles to the direction being locked. There is a separate 'slipper' of relatively hard, thin material between the screw ends and the keel. This avoids any cam action and avoids denting it.

The lever can be operated along a line within the plane of motion. This avoids skew bending and torques that will upset the adjustment while the clamp is being engaged.

The precise angle where the screws begin to clamp is easy to adjust with the slots in their ends and the split clamps on the lever. Since both screws turn, in opposite directions, the pinch is abrupt.

At times, we need a larger range of accommodation than can be easily achieved with flexures. This ball-ended link has merit for a class of devices that are switchable between the 6 DOF unlocked condition, where A can be repositioned with respect to B without restraint (within limits), and the 0 DOF situation when the end clamps are engaged.



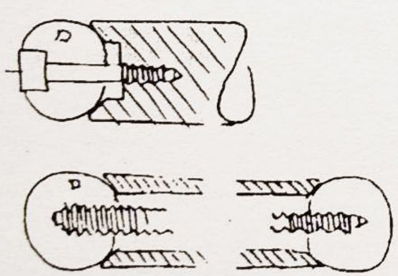
THE DUMBELL LINK

Clamping a ball in a cylindrical hole works nicely. To continue working for more than a few cycles however, Hertz must be invoked. The hole walls must be harder than the ball, else the clamping action will cause a circular dent. This makes for an unacceptable memory effect. Fabricate the clamping part out of a hardenable material.

END CLAMPS FOR DUMBELL LINK

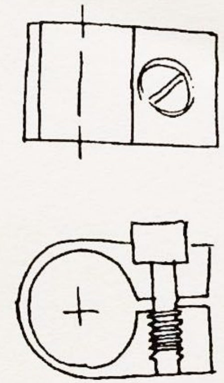
Let's not be closed-minded about ways to actually perform this clamping action. Look for 3 jaw chuck or collet action for uniform bearing so that an equal action will result in all directions.

We show some practical ways of constructing dumbbells below. A soft 18-8 cres. (300 series) ball is fine. Is this kinematic? Yes, if only one of these is used between A and B, and no other mutual constraints.



That's how it is with clamps and locks in general. *They are ideally a perfect compliment to the DOFs of the mechanism to be clamped.* If we were designing a lock (or clamp, what really's the difference?) for the two-bladed flexure, we would want it to be stiff only in the direction of the flexures. This way, its use wouldn't introduce extraneous forces.

Experience has shown that the clamp is often a source of frustration for just about everyone at the sub-micron level. If possible, beef up the adjustment parts to avoid needing locks. Qualify your design with and without and see which is best.

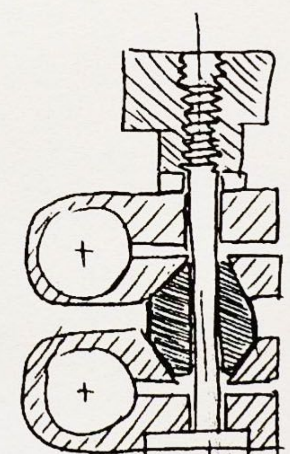


THE RING CLAMP

This ring-clamp works well when the wall is thin enough to really do some wrapping and the screw is close to the hole. Note there isn't a very good take-off point for this thing. It'll clamp a tube or rod, but then what? Often they're paired. As combined this way we call it a knuckle. The unclamped DOF count for the knuckle is only one, but because the rods can rotate about their own respective axes, a pair of knuckles can do many set-up jobs.

When the screw is tightened, often with a knob, we want the clamping action to take up smoothly in all directions, finally locking into a rigid joint. Getting this to happen requires enormous equivalent friction between the two ring clamps.

Here is a job for the taper. Ideally, a taper is a single DOF when unlocked. Its beauty lies with the behavior for various included angles. For this case, we are flirting between enough wedging action to promote high locking forces versus being able to unlock. There is no magic angle separating locking vs. not locking without specifying materials and surface finish. Something like 5 degrees is a reasonable starting point for experiment.



A KNUKLE • NOTE THE DOUBLE TAPER

The design here uses a 10 deg. pipe reamer from the hardware store as a basis. The compound rest of a lathe is then adjusted by trial and error until the double taper insert matches accurately.

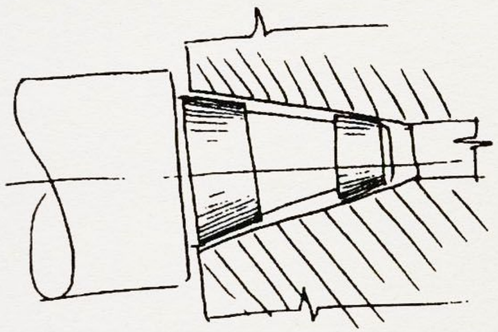
The screw can be an extension of the knob or in this case, can provide anti-rotation by flatting the sides of a screw head that goes in a slot. Only hand tightening is needed for gauging operations. Make it a point to lubricate such things if you want long life and smooth action.

We use these knuckles in our lab for many temporary holding jobs. 3/8 inch diameter stainless *tubing* is an ideal and inexpensive connecting choice. Does the knurling on clamp knobs have a real value? Yes, if and only if high manual torques are needed which should not be the case here.

A taper is not kinematic. To work well, the application must be appropriate. It's no accident that tapers are widespread in the machine tool industry. An early and still current standard is the Morse taper. This uses a small included angle and is meant to lock tightly enough to drive cutters without keys and without coming apart.

A good taper fit requires close matching over the mutual bearing areas. At least one of the parts, usually the female, is best made hardened. Relatively little foreign matter, burrs or score-marks will quickly render the Morse ineffective. Also, machine vibration caused usually by tool chatter will defeat it.

Since the premise was that the taper is locking, friction must be an essential element. Recall, we have pointed out that vibration reduces friction, sometimes when you least want it.

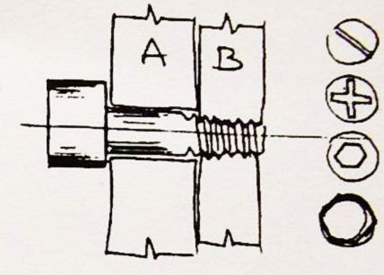


To use a taper effectively, the same shop should be responsible for *both* parts. The reason is that proper operation depends only on *differential* angle. If the parts are made to be interchangeable, then the angles must be correct on an absolute basis, requiring a male and female gauging set. Good practice relieves one or the other parts to increase the likelihood that contamination won't cause problems. Also, like the tool industry, a draw screw is a practical way of making up and breaking the tapered joint.

The taper is without peer for a rigid detachable joint between rotary parts partly because of its axisymmetric nature. In using it, you must be prepared to give up precise axial control. When axial control is required, then the taper is combined with a collet. This is a practical choice for situations where interchangeable parts of a common diameter must be included. It's not hard to buy a wide assortment of such stuff. Don't forget to consider the effect of temperature gradients and extremes on the joint's security. Some designs could incorporate a bias spring to load the taper in the make-up direction.

If you can't use all off-the-shelf standard purchased parts for a taper situation, better be prepared for a lengthy development cycle.

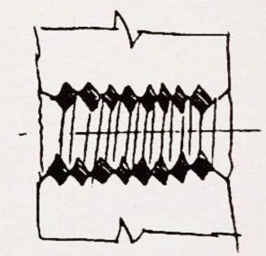
They're used for many things, but for most part, screws are simply clamps. By squeezing parts together, we expect an immobile joint. We're also looking for stability. To get satisfaction on both counts, we must do some homework.



THE MASTER CLAMP

Ideally, we want to store energy only with tension in the screw, and compression in the parts. Bending of either screw or parts is not a stable joint. Do the parts really fit together with proper bearing? When the screws go in, you want to think kinematic bearing. Coplanar pads surrounding the screw holes are nice, for both parts. The screwheads want to bear equally all around. Spherical washer sets can be helpful on both counts if desperate.

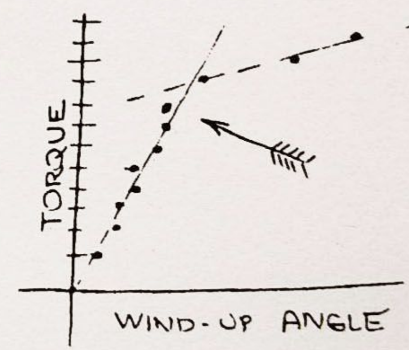
Washers are not necessarily of any help. If thick, hard and parallel, they distribute the load under the screwhead nicely, otherwise, they're probably a waste.



THE INSERT

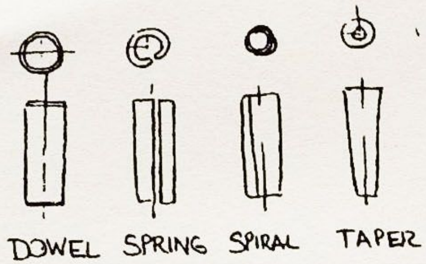
One to two diameters of thread engagement is plenty. For screws that must give high clamping forces in soft materials like aluminum, inserts are extra work, but can be had with an antivibration option that really works.

Best performance depends on a small quantity of lubrication, under the head and on the treads. They go in and out with minimal wear particulates, and they act more like ideal clamps, with tightening torque going into tensioning the screw.



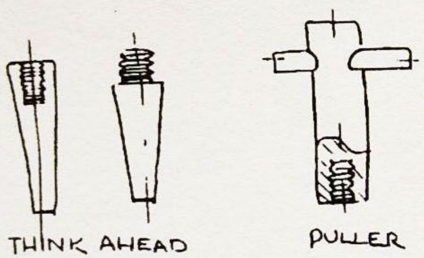
Having chosen all of the above, how much torque should you use? This you can only determine yourself with a brief experiment. Using half a dozen samples just like the production parts, prepare a graph of wind-up angle versus applied torque. Make a separate graph with all new parts for each graph. Now you can decide the answer to the torque question for yourself. The knee of the curve is the best choice for maximum performance.

Think of them as sideways clamps. Pins come in dowel, spring, spiral and taper. The dowel is most common in precision situations, and is the most demanding with respect to the details of the holes. Often a different sized hole is used in the 2 pieces being pinned for removal and replacement, which may be okay in the 10 µm region but not in submicron.



Don't ignore the pin as a machine element. A dowel can be round to a µm or less, and can be had in hard materials.

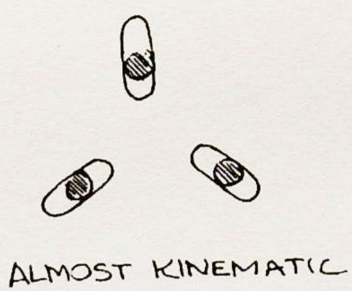
Spring and spiral pins can go into a drilled hole, and while they do some elastic averaging, it would come as a surprise to see them in our present context.



The taper pin has great merit as a removable sideways locator. A drilled hole, even made with a hand drill, is followed with a *hand operated* taper reamer. The depth is up to you.

A thru hole is always better if practical with *any* pin that is potentially removable. The metal from drilling and reaming are more easily cleaned out and dirt is less apt to become entrapped.

Only a light tap is needed to seat the taper pin securely. A pin should never have to compete with other hardware for two dimensional dominance. This means no more than one pin. Two have been used for repeated alignment, say after shipping, but only at a common temperature, and then removed after the alignment. You can change your mind by several mils by *rereaming* the tapered hole. Don't go to full depth 'till you're sure it's final.

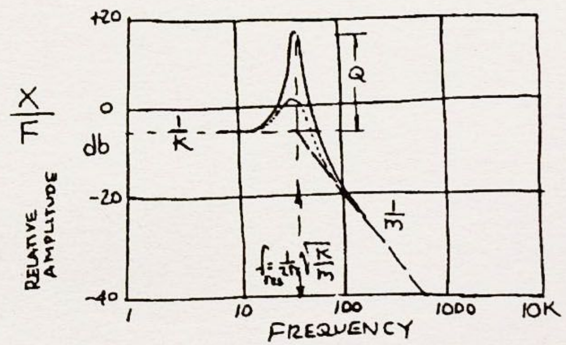


How about three dowels and three slots for X-Y-theta location? Is this not the 3 ball/3 vee again? Not quite, since a slot doesn't act like a vee in this case. Also, how to make slots with our level of precision.

Almost as a rule, we see instruments and precision mechanisms assembled on a *plate*. Many disciplines are sometimes involved: optical, electronic, computer, besides of course precision mechanics. The plate is often where they all get their stuff together. And there is usually someone around (e.g. the customer) who is partial toward light weight. The three goals affecting performance are:

> Stiffness must be high enough to be not influenced by the action of things on or off the plate.
>
> Masses and distribution shall be such that structural time constants do not influence the machine's behavior, or limit its speed.
>
> A stable reference for the interrelationship of components is expected over time, temperature, and disturbances from within and without.

If there are servo/control people looking over your shoulder, they will make certain demands for minimum resonant frequencies. Also, they may start in about high damping or low Qs. A typical assembly has many resonant frequencies, and the dominant ones (usually the lowest frequencies) are of most concern. Masses can be accurately estimated, but integrated stiffnesses are difficult. The graph below reminds us that a resonance is sensitive to the ratio of stiffness to mass.



A baseplate may be the heaviest single part of an instrument. For the same mass, surely a thinner top but with ribs would increase the stiffness. A base is especially important because it often has a master reference role to play.

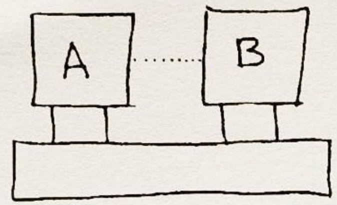
Think *boxes and tubes* rather than plates and rods. Perhaps the baseplate can be thrown out, with related things hooked directly (and kinematically) to one another. For some extra money, nearly any part can be lightweighted. This means going in with a drill and/or end mill to remove everything that isn't otherwise required where *mass can be reduced more quickly than useful stiffness*.

If we take the ever-popular slab mounting approach, access is wide open; you can add, subtract, delete, and change things around with maximum flexibility. The penalty is lots of dead weight, and poor stiffness. The trouble starts when subsystems are dependent on the plate for stability. In the event of ever present thermal gradients, there will likely be bowing. Thermal insulation reduces this effect, but at a cost of longer thermal time constants. More important may be the dynamic (read resonant) properties of the arrangement.

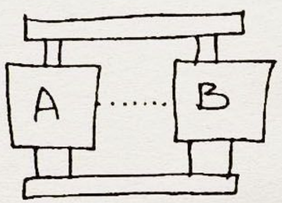
Are there relationships that are sensitive vibrational disturbances? Are there any closed-loop relationships with the structure implicitly involved as a spring/mass system?

What are the lowest resonant frequencies acceptable? What kind of damping is reasonable to assume?

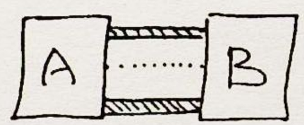
When this sort of thinking is appropriate then the box becomes attractive. Not only is the stiffness higher and mass lower but it can be made symmetrical w.r.t. its most sensitive axis. The localized tube is even harder to beat for critical connections. This is also in the popular direction of minimizing costs by minimizing weight. Temperature compensation is apt to be more successful over short, direct path lengths.



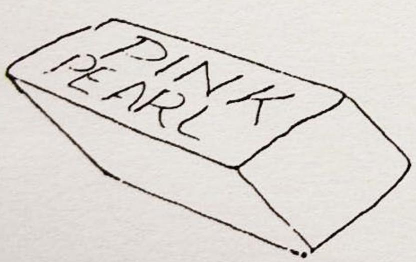
SLAB



BOX



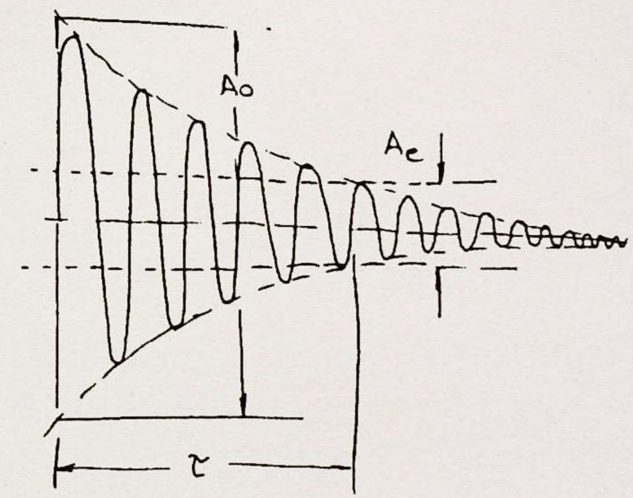
TUBE



When thinking about structural ideas remember the *Pink Pearl*. It's a good analog for thinking submicron. Just imagine your parts made of Pink Pearl and you will more easily visualize the bending, warping, sagging and flimsiness of our nanometer world. Forget about rigid. Solids may be fairly stiff, but *structures* are rubbery.

If you will have riders from other disciplines, e.g. optics, electronics, etc, then take a leadership role by enforcing a strict weight discipline.

Having maximized your stiffness to weight ratio, resonant frequencies are high. But frequency design and analysis looks at only half the picture. The other side of the coin relates to how quickly a disturbance will die out. This we characterize with the universal notion of 'Q', which is just 1/2 the reciprocal of the damping present. Somewhat paradoxically, well built structures and mechanisms tend to be characterized by their high Q because this results from low losses due to stick/slip (an amplitude dependent damping), relative motion in joints, and support interactions.

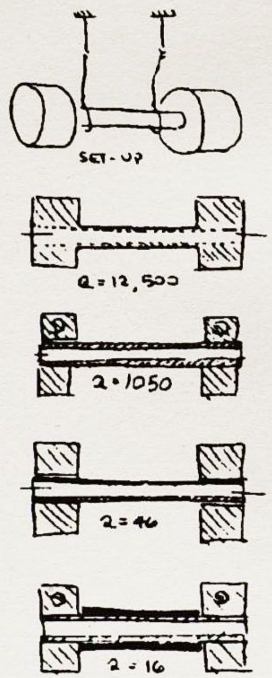


The sketch is the time domain response of a second order underdamped system with one degree of freedom, following an impulse. Real-world parts and assemblies have a myriad set of these responses, in various directions and combinations. The time domain question is how can we *damp out* vibrations as quickly as possible? For a given frequency, the length of time it takes a vibration to die out is proportional to the Q. We are looking for ways to *increase* the damping, to 'kill the Q' without introducing side effects.

We have read about the virtues of various materials as to their high damping. But in practice, the common engineering materials you want to use have almost *no* intrinsic damping! First hand experiments have taught that often-reported material damping is, in fact, measurement fixture damping! For example, a bar of 6061-T6 aluminum can exhibit a Q of 100,000. This corresponds to no useful damping whatsoever!

Don't be fooled by shaker-table responses. Even very slow-speed sweeps will not portray high Q resonances with anything like their *true* value because the length of time required to build amplitude is the same as the decay time pictured above. Sweeping through a high Q resonance may even fail to show a notable response. To build a large resonant amplitude requires that lots of energy be collected in a spring/mass system.

If we try to measure material Q, we find that there are many hidden paths for the resonant energy to 'leak out.' The method of support is crucial to this. A handy, and not too leaky a trick is to suspend the part under test by threads from above.



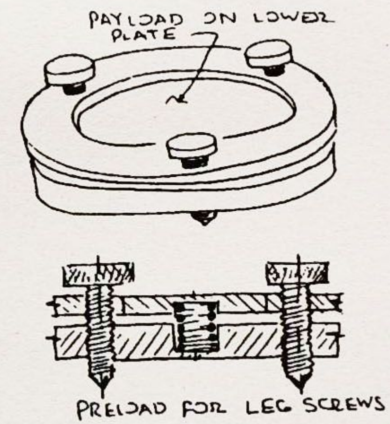
We produce a resonance by temporarily sticking a small magnet to the part with crazy glue, and exciting it with a nearby coil driven from a controlled frequency current source. An audio amplifier can be used between the frequency generator and the coil. It can take real patience to learn the art of 'pumping up' a resonance. Think of it as an oscillator. Having established the resonance, we switch off the drive and measure the length of time taken for the amplitude to die out. Our micro-gauge and a digital scope are fine for monitoring and recording the decay.

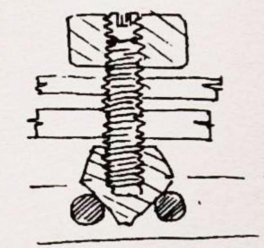
The Q is simply the length of time it takes for the amplitude to diminish to a value of $1/e$ multiplied by π and by the frequency. Or, count the number of cycles it takes for the amplitude to decay to $1/e$ and multiply by π. The EE's technique of dividing the frequency by the 1/2 power bandwidth, although correct, is hard to do in practice for a high Q situation.

To illustrate some of the factors that typically introduce damping in a practical situation, we built the series of dumbbells pictured above. The parts were all tested as described. A monolithic end-weighted tube having *no joints* gave the lowest damping. When the end weights were clamped to a separate tube, maxwell collar fashion, the damping was increased by a factor of 12:1. This is a clamped joint. When the end weights were attached with epoxy, the damping increased again by a factor of 23:1 over the clamps. Finally, an additional tube, slipped over the first, with vacuum grease in between, weights again clamped, increased the damping by a another factor of 66:1 over the clamped model.

This last technique for introducing damping can show dramatic results. Known generically as 'constrained layer damping' it depends on Newtonian shear between relatively rigid parts to absorb energy. The example illustrated here measured axial damping. Much the same results can be had in bending by adding a second piece to a part at a place where maximum shear displacements are available. Coaxial tubes don't work in bending because there's little shear. The effect is somewhat non-linear depending on other constraints between the shear parts. It is also temperature dependent because the grease has a viscosity that is temperature sensitive. Some commercial efforts have been made along similar lines.

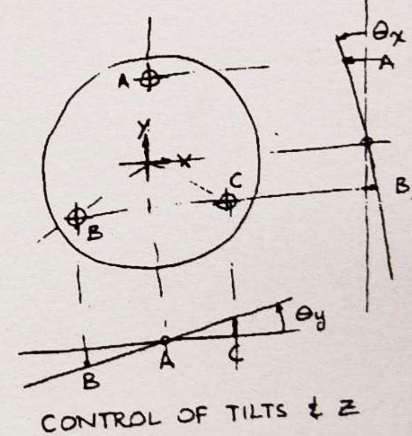
Having looked at various pieces and parts of our small-world mechanics, we will now try to put some of them together to make useful things. For the most part, we are occupied with holding and positioning components in stable, easily adjustable ways.





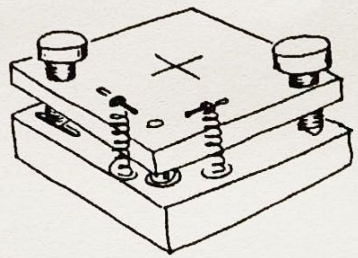
Sometimes referred to as a prism table, this device provides for an axial adjustment and two tilts. The sketch illustrates a practical way to load adjustment screws to take out wobble. We add a loading ring around the *top* of the payload support. The ring is tapped while clamped to the plate. When the three compression springs are inserted, an axial bias load is felt by the screws, which should be lubricated. By coning the screw ends, the device is a three legged stool, and can be kinematically located with the 3 vee groove plate from page 5. The addition of an extra part to the end of the screw allows a larger cone and/or a smaller screw. It should be turned after attaching to the screw.

To analyze the behavior of the adjusting screws, we imagine looking at it from the right and bottom edges. When the plate is rotated about the X and Y axes, the underlying change in screw lengths A,B, and C are obvious. From this, it is easy to write simple, linear (using small angle equivalents) relationships for the adjustments in terms of the three DOFs.

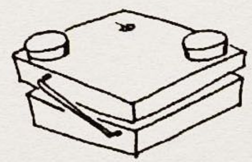


Since we have come to believe that only three legs are feasible, a choice of *where* to put them boils down to the desire for a *maximum wheelbase* in both X and Y. To avoid an unstable overturning condition, the load must be within the support triangle.

If we don't need or want the Z axis adjustment, then the classic 2 axis tilt mount may be attractive. Most people are more comfortable with adjustments at right angles (and hopefully independent of each other.) This can be found in many forms. One of the simplest, pictured here, uses a ball-in-cones for 3 constraints. With one (and only one!) of the adjustment screws in a vee, only the two tilts remain.


CLASSIC ORTHOGONAL 2 TILTS

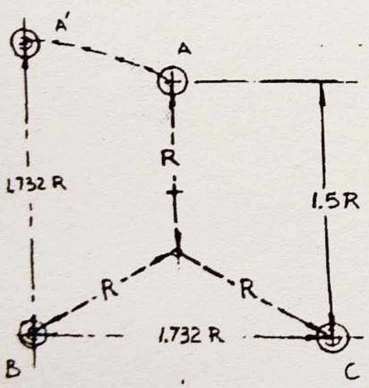
An alternative to the vee is a wire link between plates. This constraint provides antirotation and doesn't cause shearing rotation with an eccentric cone-in-vee.


WIRE LINK ON BACK EDGES PROVIDES SIMPLE ANTI-ROTATION

The diagram below attempts to highlight the differences between a *triangular* vs. a *square* layout for the ball and 2 screws.

Taking an equilateral triangle between A, B and C, the lever arm between adjustment and axis of rotation is 1.5 R. and the wheelbase is 1.732 R. If we rotate the adjustment at A to a point A', then we have a square. Now the lever arms and wheel bases are all 1.732 R.

So why even consider the triangle? Looking again, we can see that the area included within the boundaries of the ABC triangle is actually far more useful than the A'BC triangle. More important is a consideration of the relative stability of the two approaches.

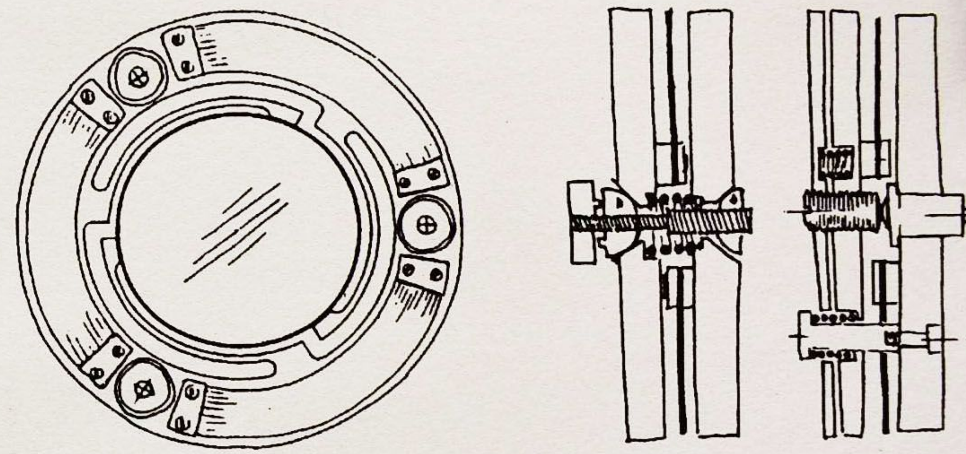


Three screws and three vees, even laid out in the square pattern; will exhibit more symmetrical properties, the equilateral with somewhat superior attractiveness due to perfect symmetry. In either case, you can designate the corner cone as a non-adjustment, even without a knob.

The springs in the upper sketch are for the purpose of keeping everything loaded so it doesn't depend on gravity to hold together.

Next, we will look at our best shot for a two or three axis mount in terms of strength, stability and precision.

The payload here is apt to be an optical part like a Fizeau reference mirror or an alignment flat. We hold the optic with a three bladed integral flexure mount (see under flexure kinematics). This allows a clear aperture with no obscuration added at the edges. Lateral position is determined by 3 crescent blades between plates (as shown earlier). For adjustment, we have a set of 3 differential coarse/fine mechanisms (also shown earlier.) These are good for .05μm setting precision, about one microradian angular settings. Adequately testing something like this is an interesting challenge in itself!

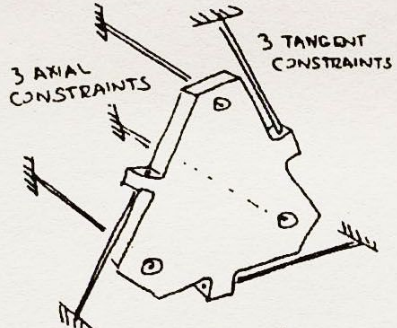


This has proven to be easy to make, with sufficient resolution for interferometry stable over long periods of time, with temperature variations of many degrees.

Careful regard for *symmetry* is our strong ally here. Also, *only localized*, closely coupled push/pull forces are involved in establishing angles and spacing. Note that unlike a set-up with two orthogonal adjusts, this arrangement can provide tilt *without* axial translation by touching up all three adjustments.

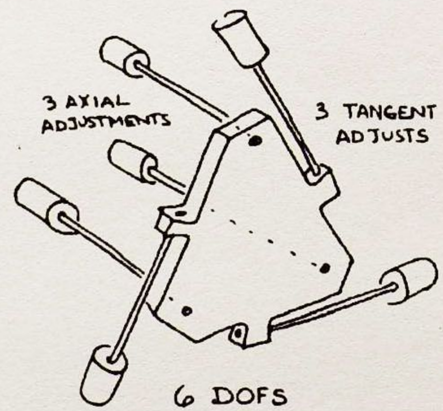
The second edge-view is a variation on the theme where the manual knobs are replaced with piezo-electric actuators for remote operation. With local feedback for each piezo, the three channels can be operated in parallel to effect a position modulation in the axial direction for phase measuring interferometry use. Note we have added a coarse manual adjust for each piezo with preloaded screws. The hold-downs take the form of die springs. This avoids the use of extension springs to clamp the plates together.

The triangular shape is symbolic of a part or component you wish to position. The sketch shows 6 kinematic connections between payload and ground. Think of these connections as a link with ball- joints on the ends. Depending on the required adjustments or motions, any combination of these links can be made fixed or adjustable. We saw earlier, examples of specific configurations chosen for their appropriateness to particular needs, whereas here we are considering the choices in a more general way.



NO DOF's

However many or few of the links are to be made adjustable, the techniques available to analyze the effect of any given actuator start with the paradoxical step of removing the adjustable link from the model. With only 5 connected links, one proceeds to analyze the remaining one DOF. This degree of freedom will, in general, have components of its path in several or all of the 3 translational and 3 rotational possibilities. Our two parallel-bladed flexure may be seen to translate mostly along the easy bending direction, but due to foreshortening of the blades, an up and down component is also present.



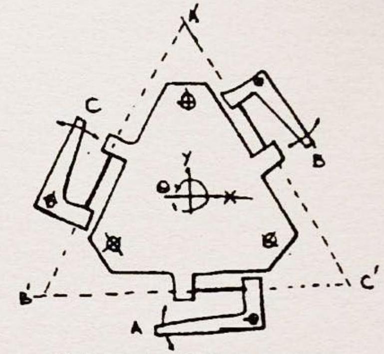
6 DOFS

A good exercise for your precision mechanics skills is to take the three axial and three tangential links pictured here and consider the displacement and rotation that would result from adjusting each link, first one at a time, and then in combinations. Note that to show a general arrangement, we have used one of several symmetric cases. The Stuart truss shown earlier is probably the most extreme symmetry case. Here, we have a symmetry in the plane with the three tangent links at 120 degrees, and symmetry in locating the plane with respect to ground, with the equilateral axial links.

Symmetry has important advantages as a candidate configuration. Aside from the improvement in stability mentioned earlier, another practical advantage lies with the relative ease of analysis and path storage, say in a computer control algorithm. It's not hard to see that an analysis of one of the axial links, as an adjustment, covers as well the other two axial links, except for a trivial coordinate system change. Similarly, a single analysis serves for the three tangential links.

The task outlined here is the compliment of the earlier case where we needed to position one plate relative to a second in the axial and tilt directions. Here we hold the axial constraints fixed, but need to position the part in the plane, call it X,Y and Θ. The upper cartoon has the symmetric tangent links of the previous page. To analyze the effect of control A, we remove the control (analytically), replace the C and D links with links to ground, and proceed to analyze the resulting four bar linkage. This will verify our intuitive guess that control at lever A causes a rotation about A'.

If this is true, then we can also believe that B causes rotation about B' etc. The simple estimates of rotations are only approximations, but are increasingly accurate for decreasingly small displacements.

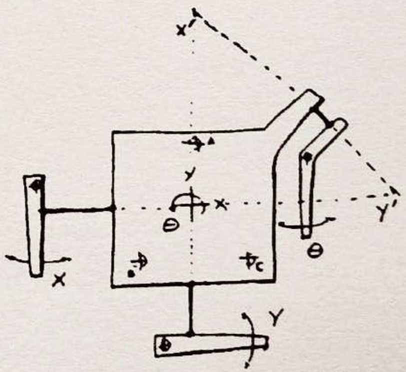


TANGENT LINK ADJUSTMENTS FOR X·Y·Θ STAGE

One could then attempt to 'invert' these control paths by parceling up desired displacement in X, Y, and Θ into changes in A,B, and C. Small displacements in X require small rotations about A' and an even smaller touch-up in the other two; an *equal* change in the three will cause pure rotation. Displacement along the X axis takes both C and B changes because neither is perpendicular to Y.

A cardboard model of proposed arrangements is helpful in gaining an insight here, as well as a rough check of your analysis.

The lower picture is an alternative (one of many) to the X, Y, Θ control problem. For a human operator, this one is sometimes thought to be an advantage. The X and Y controls approximate true motion in those directions, but are to a first order closer to rotations about the points labeled X' and Y'. The control labeled Θ is more of a pure rotation about the virtual crossing of X and Y links. The best analysis tool for this is, once again, a four bar linkage model.

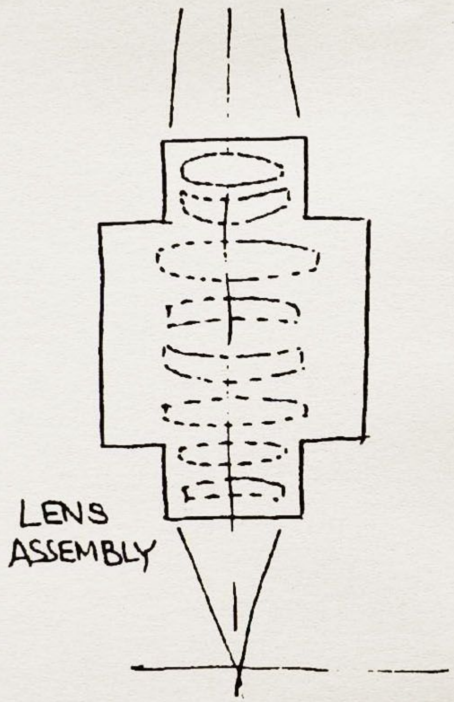


INTUITIVE X·Y Θ STAGE

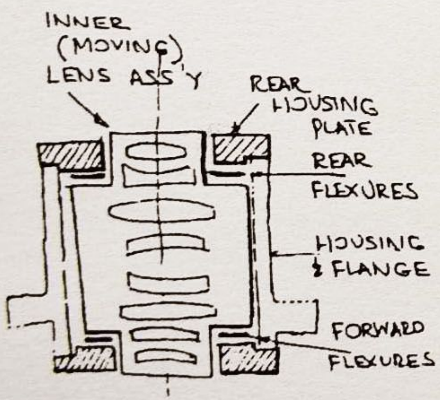
For precise adjustment, one needs a strategy for deciding on the 'steepest descent' order of adjust. For this particular case, adjusting X and Y first, perhaps with some interaction, and finally trimming the Θ control works well in practice. The principle is to leave the purest control until last.

Another practical example of precision mechanics involves the focusing mount for a high performance lens assembly. These can get as large as a foot long and 15 pounds heavy. The precision required of the best focus may be on the order of a micron, and the range of motion may be +/- a millimeter or two.

Ideally we want a *single* linear DOF with no stick/slip or lost motion. Lateral motion as a consequence of axial (focus) motion is also not acceptable here. Tilting (pitch or yaw) are also prohibited because of the requirement for uniformity of focus over the entire field. To accomplish this, we will invoke a *double* set of crescent flexures (see earlier) to supply the one DOF suspension. The lens assembly is attached to forward and rear circular housing plates with the crescents. The choice of gaps, adjacent to the blades, determines the range of motion, with builtin hard limit stops. This is precious range, so make a thoughtful choice.



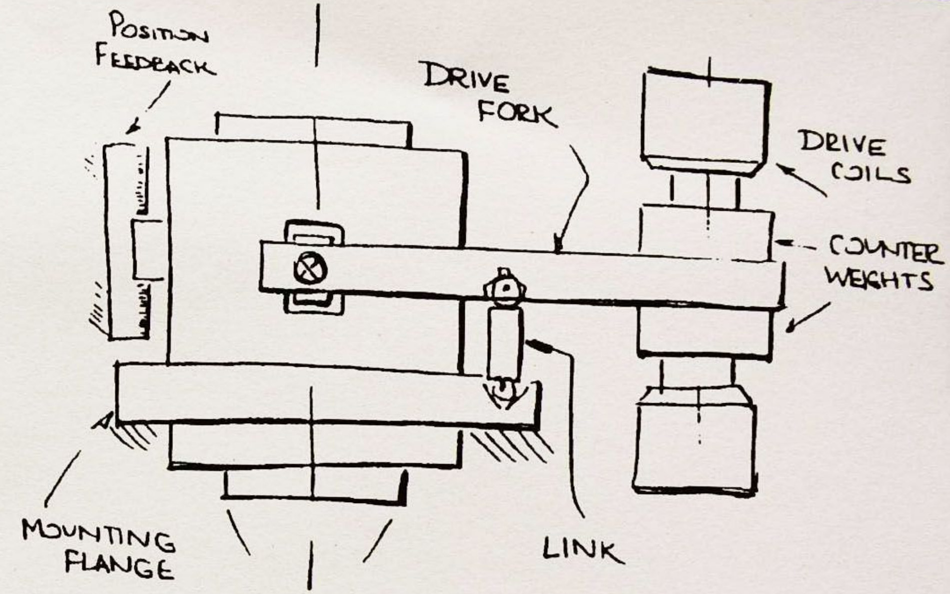
When the forward and rear housing plates are joined by a cylindrical tube and flange, the mount is essentially done. Dimensions for the blades is a trade-off between axial spring rate (want to have low) and lateral/rotational rates (want to have high).



There are several other ways to incorporate the crescents. As a guiding principle, lighter weight (for the moving part) will ease the control problem.

Note that we have provided only a mount thus far. To make it useful we need to develop an appropriate drive. The choice of flexures for the suspension was dictated by our need for freedom from stick/slip and lost motion. The same rules will also govern our choice of drive.

Ideally, the drive will act *as if* a pure axial force is applied along the optical axis. This is our next task, while preserving our kinematic relationships.



Since the flexure mount provides a single DOF, we want a drive that couples to the moving element with only a *single* constraint. The mount and drive are thus *complements* of one another; this is generally the case, equally true for drives as for clamps and locks. If the drive were to induce forces or torques other than the desired axial force, then since the lateral and rotational rates are not infinite for the mount, some error motions would be introduced.

The drive is provided here by a lever. Shaped like a fork, the ends attach to the moving part with Bendix pivots, and a single link serves the fulcrum for the lever. You may need to study this some to be convinced that, with the fork arranged this way, the lever's handle is free to move only along a predefined path (basically, an arc centered at the Bendix pivots). Counter-weights or springs can be used if gravity balance is desired for vertical orientation as shown. Counterbalancing springs are better than weights because they don't add appreciably to the net moving mass.

Finally, we have added essentials for driving the focus mount electronically. Voice coil(s), similar to a loudspeaker, have been sketched in as typical of a possible choice for a high speed, non contacting actuator. The length of the lever will determine the force/deflection relationship. For minimum power dissipation, you would like to use a minimum amount of force, but travel in this type of actuator is also limited.

An optical encoder is attached to the lens as a means of electronically measuring the actual lens position for closed-loop servo operation. Some of the basics of actuators and sensors are topics for subsequent sections. The design of a complete system depends on a nice integration of all of these parts.

Answering to the strong desire for more automated operation, remote sensing and actuation is becoming the norm in many newer designs. A typical instrument is likely to have one or more functions operating in a servo-loop for hands-off operation. An example is the electronic scale, where current through a coil is used to create a force for nulling mass-induced deflection; the current is then used to indicate mass.

In the micro world we are addressing here, there are a few special considerations worthy of our attention in the sensing and control area. Don't be scared off because of the wires. A typical sensor, even state-of-the-art types, is likely to be easy to install electrically, even though there is plenty of EE inside the box. The manufacturer will tell you everything you need to know. He is not the best expert to guide your *choice* of sensor, since he will be overly biased toward what he has for sale.

It has been said that an automatic control system is inherently limited by its feedback device. If the displacement commanded is 17.26 µm, a servo will cause corrective position displacement until the feedback element outputs a voltage or code that claims to be at 17.26 µm. But is that where it's *really* at? The answer is probably not, precisely. For one thing, there is almost never any room to put the sensor where it really belongs. In our lens focus mount, if there are any secondary motions, e.g. subtle tilts, then the encoder will be saying one thing, while the lens is doing another. Some of this can be 'calibrated out' for sure, but this is tedious and equipment and labor intensive to do in the sub micron realm. If we pay for a 0.1 µm encoder, we expect 0.1 µm accuracy! One of the key factors in a wise choice of sensor is the form factor of shape and size.

Sometimes this battle may be won by using two or more sensors in combination to derive a more accurate gauge true position. Rotary encoders are almost always used with ball bearing shafts. Bearings all have some runout, both systematic and random, thus a single 'read' station is going to suffer errors even given a perfect, noise-free encoder disc. This problem is significantly reduced by using several (usually two or four) read stations. In our lens mount, we could use two linear encoders and average out tilt, or, using three encoders we could average for the focus, and derive signals for residual tilt correction, if required. This sounds expensive, but in principle is valid.

Most, if not all, of our lessons for 'degrees of freedom' apply analogously to measuring. If there is strictly one DOF, it probably isn't an exact straight line or a perfectly circular arc. A designer is advised to consider all potential geometric error sources on the measurement. In the not uncommon situation where more than one DOF must be controlled, a like number of sensors will be needed, although not necessarily in a one-for-one relationship.

In the case of an X-Y-Θ stage, it's not obvious where one would locate sensors (at least 3) to provide accurate remote sensing. If this turns out to be your case, then all is not lost! Often you will find adequately proximate locations for sensors, but that don't output your desired coordinates. This can be handled by analyzing what the geometry will output at the sensors for stage displacement in X, Y, and Θ, taken one at a time. If the range of motions is small, very often the resulting sensor outputs can be combined algebraically to give the desired signals. You may find that the sensor(s) can't be readily integrated in a design as an afterthought.

To develop some feeling for small displacement gauges, we will explore the more important properties, with comments on *why* they are significant and how to measure them. Certainly manufacturers' data is useful, but remember, they want to sell you something, and in general, you will find that not all parameters are well characterized.

Being practical minded, we will concentrate on the more commonly used gauges that can be purchased as complete components, later discussing some you can make yourself in volume at very low cost. Proximity types are attractive because they tend to be insensitive to all but changes in a gap between sensor and target. Of these, capacitive, eddy-current, and photo reflective are the most common. When specifying a gauge, we will want to take into account at least the following properties:

RANGE

If the requirement is for 100 µm total displacement, we want a device with 4 1/2 or 5 mil range. If a 10 or 25 mil device is used, it will probably be larger and have a less favorable signal-to-noise ratio over *our* desired range. Operating range is often over a gap with a rather soft upper bound. You are going to need some calibration equipment, but don't rush into buying the fixtures of a manufacturer until you are confident of what is really needed. Look for the minimum gap recommended. This is often zero. Better not try operating down to zero gap because a crunch is practically guaranteed sooner or later, disturbing the relationship between your moving part and the sensor.

TARGET

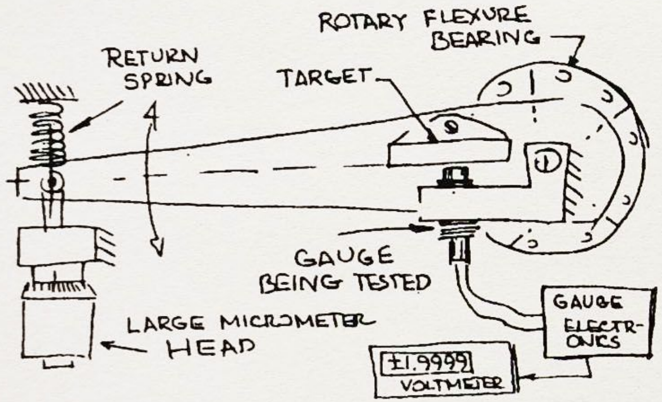
Capacitive gauges require a conducting, grounded target; the material isn't significant. For the eddy-current gauge, most any metal will work, but you need to read manufacturer's recommendations for the metal choice vs. the particular candidate gauge, as calibration *will* be affected. Aluminum tape has been found to work well. The photo-reflective sensor is target reflectivity sensitive, and must therefore be calibrated with this in mind. A stick-on white paper label works well enough. Target *size* is also a consideration. The photo-reflective sensor uses the smallest spot. If the target will move laterally, the eddy-current gauge may react to variations in target *alloy* variations, although we haven't seen this with aluminum tape.

When choosing the location for a gauge, avoid any situation where a sensor cable will be subject to twisting or bending motions. The eddy-current gauge in particular will generate a phantom signal with cable flexing. If possible, you will be happiest with the sensor stationary. If cable motion is your desire, better plan an experiment before committing to a particular gauge and configuration.

LINEARITY

This term derives from the expression, $Y=mX+b$, where m is the gauge factor, usually in volts per unit displacement, X is the gap or displacement, and b is the voltage when the displacement (or gap) is zero. Linearity is a measure of how closely a particular gauge comes to this characterization. You, or your system, will be using a voltage to infer displacement. If you know the gauge factor accurately, and the device is stable and linear, then the inference can be quite accurate. Determining the gauge factor and linearity is important to accurate sensing, thus we will look at the methodology for doing a calibration next.

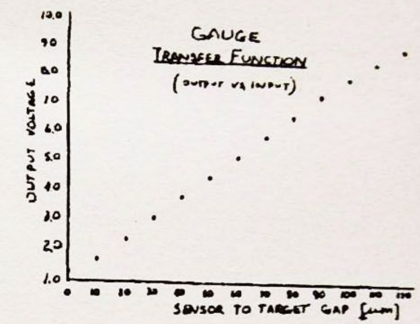
Here we have a typical gauge attached to a calibration fixture. We are taking advantage of the fact that proximity sensors tend to average across their working faces, thus small angles don't influence the result. This allows us to use a lever reduction between the micrometer head and the sensor. In order to insure precise rotation we have incorporated a flexure bearing (see paddlewheel bearing earlier) to create the pivot. With the micrometer head at a 13 inch working radius and the gauge at 1.3 inches, we have a 10:1 reduction between drive and sense. This also divides micrometer head errors by the same ratio.



With the micrometer head backed off to its zero position, the gauge and target are adjusted for *square* contact. For precision work it is best to use a low power microscope to set the micrometer head; this will eliminate parallax-induced uncertainties in settings. A quality micrometer head is specified at better than 50 microinch (or 1 μm) accuracy, which translates to 5 microinch or 0.1 μm at the gauge. In our example we are going to calibrate a gauge over a range of 5 mils (127 μm), thus we would expect the micrometer head to contribute less than 1 part in 1000 (0.1%) to the combined errors measured. In actual practice, we have come to expect considerably less.

Manufacturers want to sell you a micrometer head with targets mounted to its spindle. This will put the entire head error into your calibration, and with no loading to the micrometer screw, it will not be as likely to give its best performance. The use of stepped spacers or ceramic gauge blocks is (at best) a quick and dirty check for gross malfunction of a gauge, but will never yield a truly accurate calibration. Yes, we could imagine using a servo-driven air slide and laser interferometer for a high class calibrator, but this begins to sound expensive and is probably not warranted except perhaps for a sensor OEM.

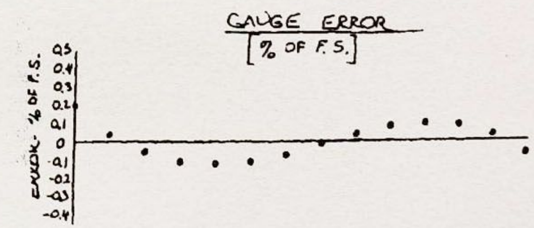
Having set up our calibrator, we need to find a suitable power supply and read-out, probably a 4 1/2 digit voltmeter. A gauge that outputs a *bipolar* +/- voltage, say +/- 10 volts is preferable because it utilizes the full span of a typical meter better. And, when taking readings, it's best to avoid changing ranges on the meter.



The points plotted on the graph here are typical of what you would expect from a linear sensor. Note there is a subtle rounding at the bottom and top ends (known in the trade as toe and shoulder) which would get worse quickly if we attempted to extend the operating range beyond what's shown.

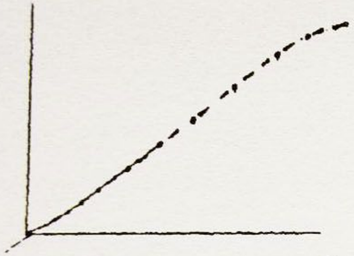
A straight plot like this doesn't show you very much, since we're apt to be looking for fractions of a % deviation from straight line behavior. Instead, we invoke the least squares (linear regression) program built into many engineering calculators to produce the 'm' and 'b' implied by our data. We then compute for each data point, the ideal straight line voltage. The graph below plots the *difference* between this *ideal* voltage and the *actual* measured voltage divided by the range (in the example shown the output range is about 9.5 volts) and multiplied by 100 to express the error as a percent of full scale.

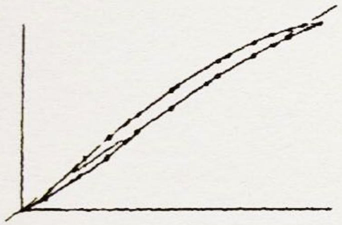
When these errors are plotted, we have the much more meaningful curve shown below. This is a classical error curve; it tells us quite a lot about the gauge. It also says something about the precision of our data taking! Notice that the errors form a smooth, low order polynomial. With less care in setting the input or stretching the measuring time over a long period, considerable noise would be evident.



A well executed calibration run really does look like this! If yours doesn't seem to be smooth, try to do it faster with two people, one to set the input and another to record the meter output values. Run thru the procedure 3 or 4 times 'till it can be done quickly. Less time means tighter data. We use a 488 bus from the voltmeter to a computer so that a foot switch causes each reading to be recorded. The error curve shows a cubic residual and suggests that a cubic rather than a linear fit would be more accurate. This is generally true, if overall stability supports the sophistication.

One other aspect of linearity not covered by the discussion thus far is hysteresis. Many sensors (and most test fixtures) exhibit a memory effect in that its output depends on prior position. This is clearly displayed by doing the calibration, data reduction, and error curve as we did it, except we would include data for the sequence from maximum input back down to zero again.





Analyzing both up and down readings will yield a transfer function that is single valued, as the upper graph shows if the device is without hysteresis. This is what you are looking for.

The lower graph shows different output for the same input, depending on the direction data are being taken. Actually, for a new device you are exploring this is an important test. Even more sensitive will be the resulting error curve, since it magnifies nonlinearities.

If your data does show hysteresis, the next question is why? Ask the manufacturer what should be expected. This is rarely specified, though important. If the device is not supposed to have any memory effect, then maybe it's in your test equipment.

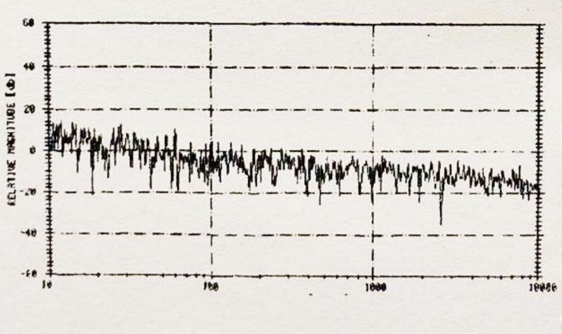
Having digested range and linearity, we will go on with our menu of sensor characteristics that deal with dynamic properties.

FREQUENCY RESPONSE

This is especially important for feedback in a closed-loop situation, along with its partner, phase response. Actually, we mean 'amplitude and phase vs. frequency.' Capacitor, eddy-current, and photo-proximity gauges are all capable of 10 Kilohertz response. In a loop, the useful range of frequencies is apt to be limited by phase response; 10 or 15 degrees of phase shift is precious to the loop designer. In general however, excess frequency response (beyond what you need for your application) becomes a liability.

NOISE

This is the random voltage that outputs superimposed on the signal that you want. For a given sensor, the noise tends to rise as the square root of the bandwidth so that by using a cut-off filter you can have lower net noise for the same (low frequency) signal. Noise, frequency response and resolution are closely tied together. This is a plot of noise as evidenced in the frequency domain. Note log scale.

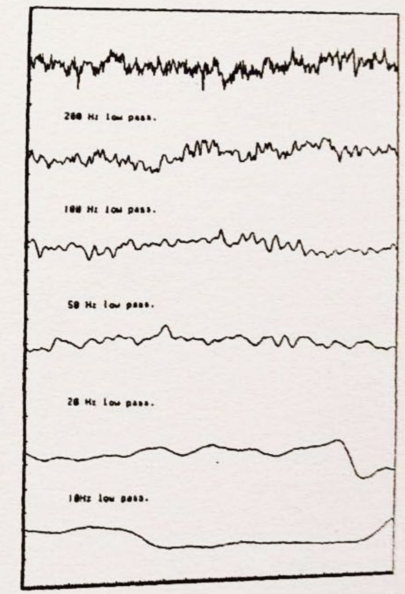


NOISE EQUIVALENT DISPLACEMENT

For a given sensor (including amplifiers, filters, etc.) there will be a noise output (with no input displacement). When the noise is measured, conventionally as either RMS or Peak-to-Peak, we can use our gauge factor, or its inverse, which will have units of displacement vs. voltage. This, times the measured noise, is the system's *noise equivalent displacement*. The sensor is telling you that there is this noisy displacement even when there is actually none. This is one measure of the system's resolution. A real displacement on the same order as the noise is contaminated with this noise fluctuation. Put in another way, this is the signal size where the signal to noise ratio is 1.

DYNAMIC RANGE

This is simply the maximum linear signal divided by the noise equivalent displacement. Like signal-to-noise, dynamic range increases with lower frequency cut-off. For a low frequency system, dynamic range can go to 80 or 100 db (10,000 or 100,000:1). For a system with a wide bandwidth, 60 db (1000:1) is more like what you can expect.



This sequence of wiggly lines is the *time domain* way of looking at noise. Note that the character and amplitude of the noise changes as we introduce lower and lower cut-off frequency with a separate low pass filter. Compare this with the *frequency domain* graph on the previous page.

Since we nearly always are looking for a position indication from a gauge, all of the measurements we have been making are DC coupled. An exception is a velocity feedback device for resonant scanners or tuning forks. A loop can be built to maintain constant amplitude by keeping the velocity constant. This is because the losses in a resonant device are proportional to the velocity. A drive coil current is used to maintain constant velocity. This can be a simple untuned amplifier where the mechanical oscillator will choose its own frequency if the drive is taken from the rotor's velocity.

STABILITY

For a fixed displacement, how much will the output wander around as the temperature is changed? And how much will it change just due to time? These effects tend to be superimposed in a real system. When manufacturers quote stability vs. temperature, do they mean just the sensor itself or does it include the supporting electronics as well? Actually, since we are implicitly assuming a V=mX+b sensor, both 'm' and 'b' will be effected by time and temperature. Build a test fixture that holds the gauge immobile at mid range, using the same materials intended for your application, and measure stability for sensor and electronics with separate temperature environments.

David H. Kittell was born in 1935, in Haverhill, MA. He attended his much-loved Proctor Academy in Andover, NH, during his high school years, Dartmouth College for two years, New York School of Photography, and finally the University of Connecticut, graduating with high honors from the School of Engineering.

His early career as a photographer brought him into contact with the world of engineering at Barnes Engineering and Time, Inc., after which he started his consulting firm Sage Systems. Over the long years of his career, he worked mostly for PerkinElmer on NASA projects. His work included the design of PerkinElmer's computer-controlled polisher that polished the famous Hubble Telescope mirror. He developed the world's first deformable mirrors, first for the SDI "Star Wars" program and then for the world of celestial observatories. (Nearly all of the world's big telescopes now feature a deformable mirror as a core component.) His crowning achievement in the field of precision mechanics was the world's first fully operational Fabry-Pérot interferometer which required single-angstrom-level positional stability of a pair of six-inch diameter mirrors for periods as long as twelve hours.

Over the years he developed the discipline of Precision Mechanics, which he first taught at the PerkinElmer Technical Institute. He later taught the course in many formats to innovators at companies including Bell Labs, The Society of Photographic Instrumentation Engineers (SPIE), the American Society of Precision Engineering (ASPE), Lawrence Livermore National Labs, and many others. This book was the end result of this evolution.

David was an avid animal lover, a patron of the musical arts, and a most inspiring father.

Please enjoy this book knowing it is a work of devotion and love.

—Jake Kittell
Morrisville, VT
October 2017

PRECISION MECHANICS COURSE OFFERED

OVERVIEW — An on-site, tailored, two-to-three-day course for five to several hundred participants. Highly practical with many hands-on demonstrations throughout, this course will be valued by anyone who spends considerable time around machines. Participants will each receive a copy of *Precision Mechanics*, the "little blue book" by David Kittell.

OBJECTIVE — Provide participants with practical, applicable techniques and approaches for the engineering design of instruments and precision devices.

WHO SHOULD ATTEND — Engineers, scientists, technicians, machinists; anyone who needs to create instruments, mechanisms, mounts, fixtures, adjusting, or supporting devices with precision performance requirements.

DESCRIPTION — The course begins with the concept of kinematic design and how to create practical elements for kinematic mounts. It goes on to discuss flexures for small displacements and rotations; wire flexures for 5 DOF coupling; methods for precision adjustments; holding, clamping, and locking methods; considerations in supporting structures—their stiffness, vibration, and damping; stick-slip; Hertzian contact stresses; stability over time and temperature; and many examples that show how these methods can be integrated to achieve a desired result. Model shop and experimental lab techniques are often referenced. Intuitive understanding of fundamental principles is emphasized. Measurement components and application to the sub-micron world are discussed.

INSTRUCTOR BIO — Jake Kittell is principal engineer at Happiness Tech LLC and research engineer at the University of Vermont. He holds an engineering degree from the University of Connecticut and has built precision machines and robots for over 25 years.

CONTACT JAKE KITTELL FOR MORE INFORMATION

jake@happinesstech.biz

Precision Mechanics

D. Kittell