FUN*da*MENTALS of Design

Topic 3

FUN*da****MENTAL*** **Principles**

FUNdaMENTAL Principles

Imagine the feeling you get when you participate in an activity in which you RULE! When you MASTER the *FUNdaMENTALs* of design, you get the same feeling, continuously! This entire book is really about helping you to learn to master the *FUNdaMENTALs!*

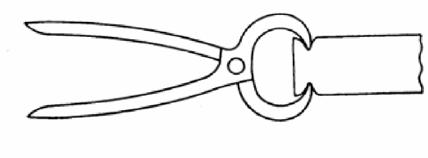
Long before any detailed design engineering is begun, an engineer has to have a vision of the machine concept in mind. The creation of the vision occurs in the engineer's bio neural net; and the creativity and efficiency of the bio neural net is affected by the depth of understanding of fundamental principles. These fundamental principles are best learned by experimentation, both physical and analytical!

This chapter focuses on the philosophy of the physics of the design of machines. With a deep knowledge of these fundamental principles, one can rapidly generate strategies and concepts with the greatest viability. Then, when it comes to the detailed engineering phase of the design process, analysis will be more fruitful, and you will encounter far fewer dead-ends that require you to start over.

Furthermore, with a deep understanding of fundamental principles, one can more critically evaluate other machines and components; hence others will seek you out for design reviews, which will further deepen your knowledge of the fundamentals and increase your knowledge of what is new and exciting!

Hence in many respects, this chapter is the foundation upon which all other chapters are built. This chapter should be read and re-read many times, until every principle and every picture is firmly ingrained in your mind. Furthermore, the blue thought exercises at the end of each page should be diligently undertaken, for they can help to identify problems (opportunities!) early-on in the design process, which is a key to minimizing cost and pain.

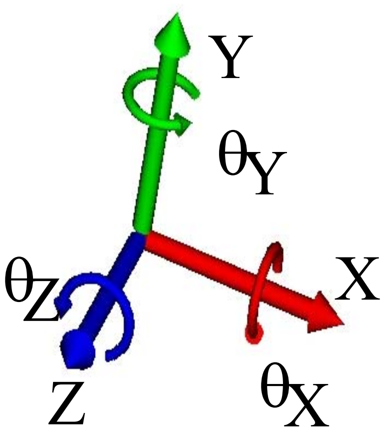
The passion with which you pursue a design is often affected by the probability of success that you feel is possible. A deep understanding of the fundamentals is as powerful a design aphrodisiac as one can have! Go ahead, make your day, indulge your mind!

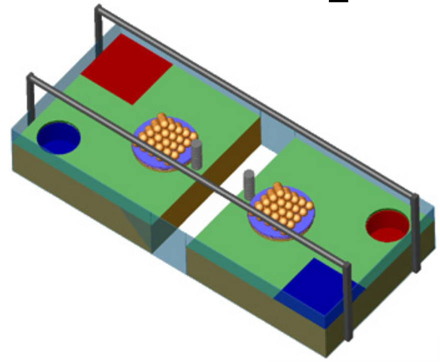


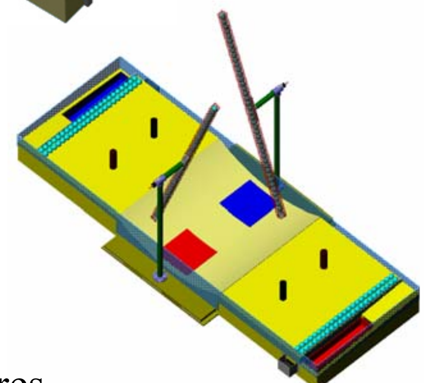
Topic 3
FUNdaMENTAL Principles

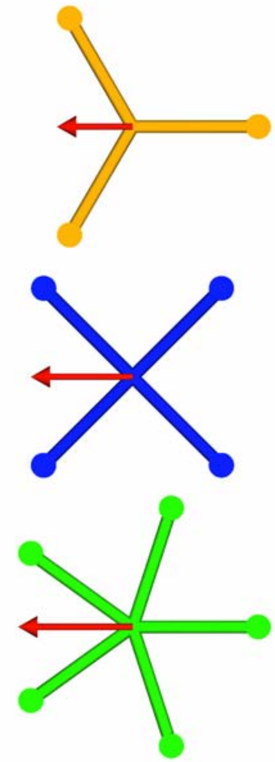
Topics

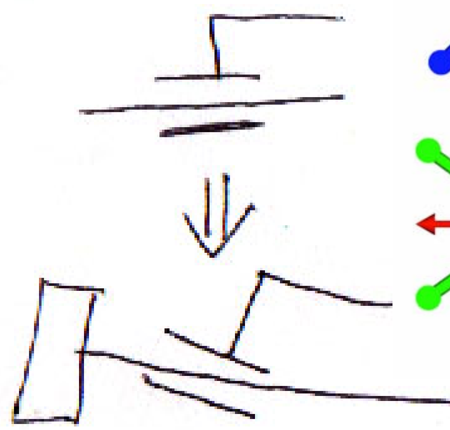
- Occam's Razor
- Newton's Laws
- Conservation of Energy
- Saint-Venant's Principle
- Golden Rectangle
- Abbe's Principle
- Maxwell & Reciprocity
- Self-Principles
- Stability
- Symmetry
- Parallel Axis Theorem
- Accuracy, Repeatability, Resolution
- Sensitive Directions & Reference Features
- Structural Loops
- Preload
- Centers of Action
- Exact Constraint Design
- Elastically Averaged Design
- Stick Figures











Occam's Razor

All too common an engineer laments "it's too complex", or "I should have made it simpler". In hindsight, we often have 20/20 vision; so why not minimize complexity from the beginning? Complexity is to be minimized in both design and in manufacturing, which are often intertwined; hence in order to be successful, a careful systems approach is required. Designers should "Keep It Super Simple" (KISS) and "Make It Super Simple" (MISS)!

The first opportunity to keep things simple is with the overall design ***strategy***. If a less complex method is created in the first place, the rest of the design will likely be less complex. When using FRDPARRC tables, one of the key risk assessment criteria is complexity. For example, in the robot design contest *The MIT and the Pendulum*, consider how to get the balls out of the pendulum. One could design a machine to climb the pendulum and engage the axle to turn the pendulum, or one could have a vehicle that drives forward fast enough to knock the pendulum to a height that dumps the balls from inside it (see page 2-20).

The second opportunity to keep things simple is with the design ***concept***. Once again, the FRDPARRC table can be of great help as it asks you to not only consider risk, but references. You are less likely to select a complex solution if you can see that your opponents have a simple effective solution. If a design parameter (idea) has large risk, including "too many parts" then you should think of simpler ideas that can accomplish the same functional requirement. The ***concept*** phase is also where disruptive technologies can manifest themselves to your advantage. An example is the use of linear electric motors and air bearings in a simple shape to create the Axtrusion™ linear motion axis (See page 1-14).

The third opportunity to keep things simple is with the formation of ***modules***. The ***modules'*** functional requirements come from the design parameters of the ***concept's*** FRDPARRC Table and this offers a another chance to reduce complexity. If a ***module*** becomes very complex, there is still time to go back and try to simplify the ***strategy*** or ***concept***. This is also a good opportunity to specifically consider manufacturing issues because the ***modules*** will have to be assembled and then interfaced to each other. This is where *Design for Assembly* (DFA) starts to enter the picture.

The fourth opportunity to keep things simple and also further address manufacturing, is with the design and selection of ***components***. This is often the make-or-break phase of KISS & MISS as it is at the ***component*** level where the design details get done. A first knee-jerk reaction is to simply reduce the number of ***components***; however this can sometimes cause more problems than it solves; hence it is critical to keep not only the functional requirements of the design in mind, but also the implied functional requirements of the ***components*** and their operation.

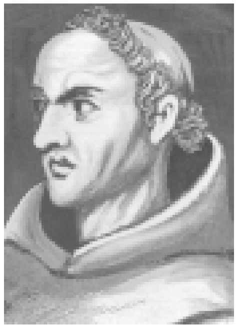
A common example involves designing systems with bearings, where improper constraint is often caused by a lack of attention to detail. It may seem like a money saving proposition to reduce the number of bolts that hold a bearing preload cover in place; however, reducing the number of bolts can make the deformation zones (strain cones) under the bolt heads no longer overlap. This causes the bearing to be subjected to point loads, as opposed to a quasi isostatic load, and then the bearing goes bump-bump-bump-bump and soon fails. The stiffness of a fully bolted structure where the strain cones overlap can be modeled as if the components where welded together. Using just enough bolts to withstand applied forces, on the other hand, can lead to inadequate stiffness.

Sometimes reducing the number of ***components*** leads to specification of a few highly complex ***components***. Thus the goal is to minimize the number of ***components***, and balance ***component*** complexity with the number and complexity of manufacturing processes and the quantity to be manufactured. In large quantities, such as in automotive applications, minimizing material mass and number of ***components*** often outweighs any perceived manufacturing complexity, because process machines and tooling are quickly amortized. Indeed, there are formal methods of designing for assembly, which often focus on reducing part count and directions of approach during assembly. Snap fits are a classic method for reducing complexity. However, one has to balance tooling costs with production volume, and the ability to evolve the product.

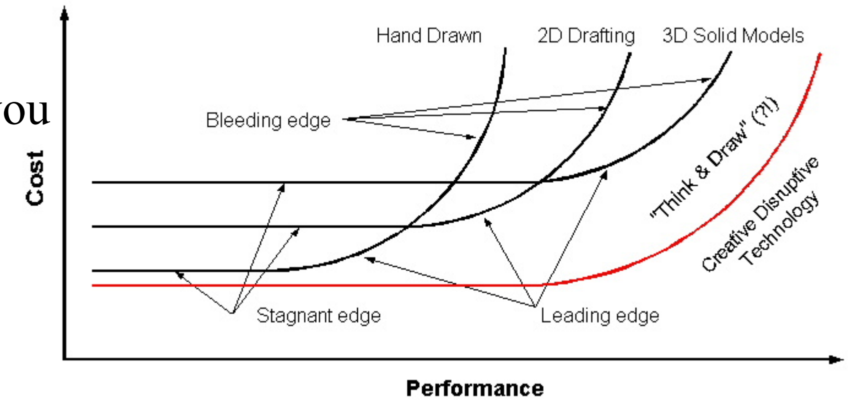
Review your various robot designs according to the above thoughts, evaluate their complexity. Look at every region where there is a joint and ask yourself how it will be fastened, and could instead the part be made from a piece of folded sheet metal or machined from a monolithic block of material? Apply *Maudslay's Maxims* as discussed on page 1-3.

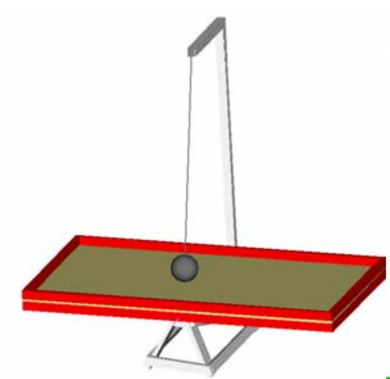
Occam’s Razor

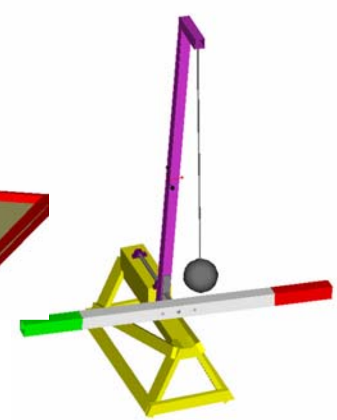
- William of Occam (Ockham) (1284-1347) was an English philosopher and theologian
  - Ockham stressed the Aristotelian principle that *entities must not be multiplied beyond what is necessary (see Maudslay’s maxims on page 1-4)*
    - “The medieval rule of parsimony, or principle of economy, frequently used by Ockham came to be known as **Ockham's razor**. The rule, which said that *plurality should not be assumed without necessity* (or, in modern English, *keep it simple, stupid*), was used to eliminate many pseudo-explanatory entities” (http://wotug.ukc.ac.uk/parallel/www/occam/occam-bio.html)
    - ***A problem should be stated in its most basic and simplest terms***
    - ***The simplest theory that fits the facts of a problem is the one that should be selected***
    - ***Limit Analysis can be used to check ideas***

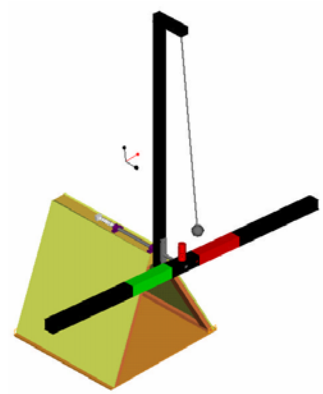


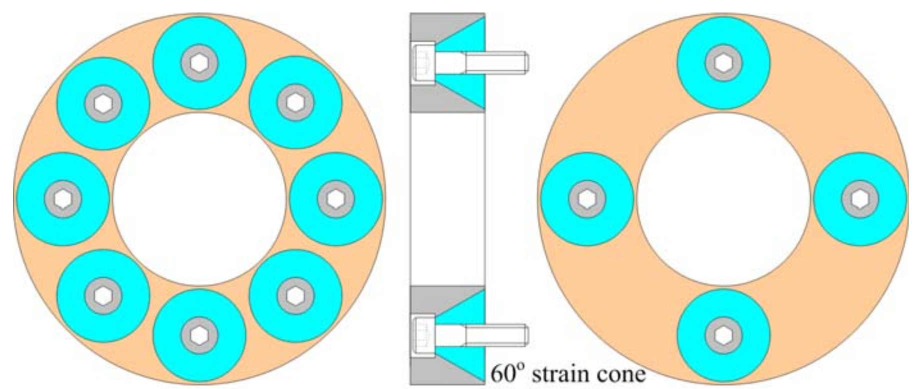
- Use fundamental principles as catalysts to help you
  - Keep It Super Simple (KISS)
  - Make It Super Simple (MISS)
  - *“Silicon is cheaper than cast iron”* (Don Blomquist)











Newton's Laws

Newton's Laws provided the foundation for the study of the mechanics of solids and fluids and catalyzed the industrial revolution: This led to the formation of the engineering profession which enables us to engineer, as oppose to develop by trial & error, machines. In addition, Newton's systematic methods of discovery and mathematical modeling form the foundation of the scientific method, which is the basis for a deterministic design process for developing machines.

Newton's ***First Law*** sets the stage for the motion of objects: *Every body persists in its state of rest or of uniform motion in a straight line unless it is compelled to change that state by forces imposed on it.* This law is what we use when we say the sum of forces must be zero, and the sum of moments about a point must be zero: *Force and moment equilibrium* must exist.

Newton's ***Second Law*** is a generalized version of the first law: *The acceleration of a body is directly proportional to the resultant force acting on it and parallel in direction to this force and that the acceleration, for a given force, is inversely proportional to the mass of the body.* The same is true for rotary motion systems where a force $\boldsymbol{F}$ is replaced by a torque Γ, mass m by moment of inertia J, and linear acceleration $\boldsymbol{a}$ by angular acceleration α:[1]

$$\mathbf{F} = m\mathbf{a} \qquad \mathbf{\Gamma} = J\boldsymbol{\alpha}$$

The Second Law gives rise to the differential expressions of motion:

$$\frac{dx}{dt} = V \quad \Rightarrow x(t) = x_{t=0} + \int_0^t V dt$$
$$\frac{dV}{dt} = a \quad \Rightarrow V(t) = V_{t=0} + \int_0^t a dt$$

Newton's ***Third Law*** states: *To every action there is always opposed an equal reaction or, the mutual actions of two bodies upon each other are always equal and directed to contrary parts.* Newton's third law directly leads to the principle of ***conservation of linear and angular momentum*****:** *When the resultant external force acting on a system is zero, the total vector momentum of the system remains constant.* This is true for both linear, $\boldsymbol{P}$, and angular, $\boldsymbol{L}$, momentum:

$$\frac{d\mathbf{P}}{dt} = 0 \text{ or } m\mathbf{v} = \text{constant} \quad \text{when no external forces are applied to the system}$$
$$\frac{d\mathbf{L}}{dt} = 0 \text{ or } J\boldsymbol{\omega} = \text{constant} \quad \text{when no external torques are applied to the system}$$

If an ice-skater (assume zero friction between her and the ground) where to shoot a bow and arrow in the direction in which her skates were pointing, the product of her mass and resulting rearward velocity would equal the product of the mass of the arrow and its forward velocity. Even though energy was expended in the process, it could not change the momentum of the system because it did not impart any forces to the environment.

How does a cat always right itself when it falls, or a diver or gymnast twist in the air? The lifeform imparts a small torque or moment[2] as it is launched and thus starts its trajectory with some angular velocity which it then modifies by changing its moment of inertia. For example, a diver's arms are swinging as he leaves the board, and by using internal work to move the arms and legs, the tuck, the diver can greatly decrease his moment of inertia and temporarily increase his angular velocity. Since the diver used his muscles to do work on his body, energy was not conserved![3] One part of the body can move with respect to the other, but angular momentum will always be conserved.

Under what circumstances does your ***strategy*** or ***concept*** require the use of conservation of momentum or the conservation of energy? How can you use these basic principles to help you quickly sort through ideas? Sketch and use *Free Body Diagrams* to help you do force and moment accounting.

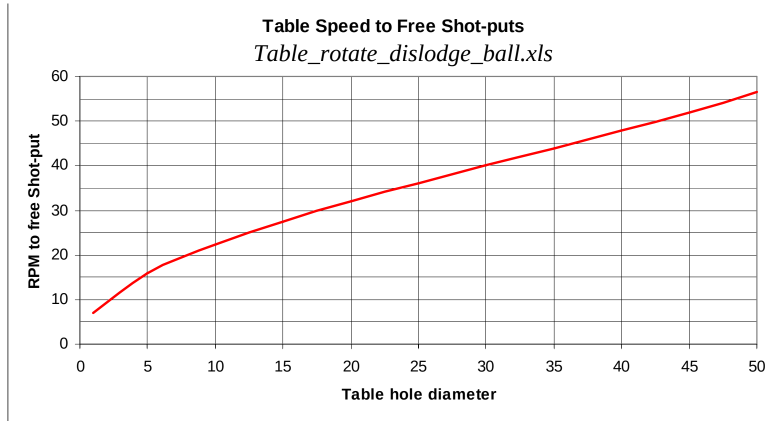
---

1. Boldface variables indicates variables that are vectors: they have magnitude and direction.

2. A *torque* is a twisting "force" that causes pure rotation. A *moment* is also a "twisting force" but it can also cause linear displacement. For example, a T-handle wrench allows you to pull with one hand and push with the other, so you can turn a bolt without the wrench being forced sideways and off the bolt. Just pull on one handle of the T-handle and the wrench is likely to come off the bolt. Do not get too hung up on the term "moment" verses "torque" as they have the same general effect and the same units of N-m.
3. See R.L. Page "The Mechanics of Swimming and Diving" The Physics Teacher, Feb. 1976.

Newton's Laws

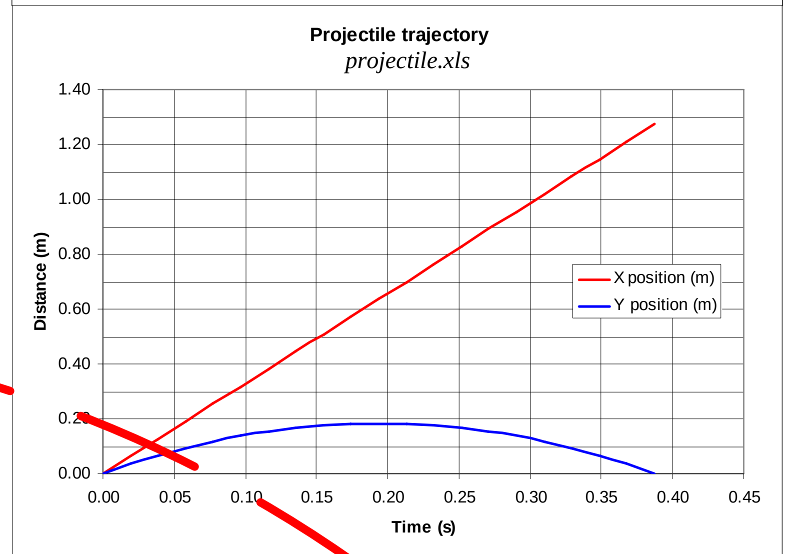
- 1st, 2nd, & 3rd "Laws" are invaluable design catalysts that can help launch many an idea!
  - (The only real "law", perhaps, is 300,000 km/second!)
- Conservation of linear momentum
  - If no force is applied, then momentum is constant
- Conservation of angular momentum
  - If no torque is applied to a body about an axis, angular momentum is constant about that axis
    - A force coincident with an axis does not apply torque about that axis



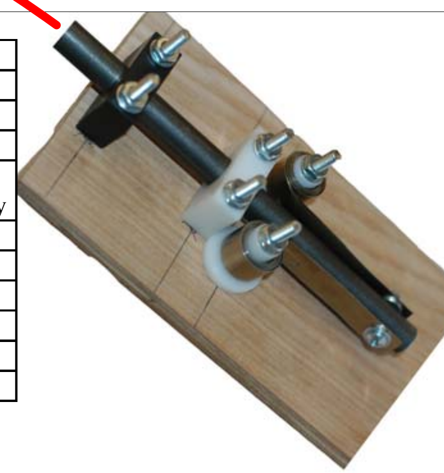
See Projectile_motion.xls

Sir Isaac Newton (1642 - 1727)

Isaac probably would have LOVED snowboarding!



Power_budget_estimate.xls										
Power & energy budget for individual moves, total (S) for simultaneous moves, and cumulative										
Last modified 9/01/03 by Alex Slocum										
Enters numbers in **BOLD,** Results in RED						Power (Watts)			Energy (N-m)	
Axis	Move #	Force (N)	Velocity (m/s)	Distance (m)	Efficiency, net system	Move	Battery dissipation	Σ power for move #	Energy for move	Σ Energy
Drive to pucks	**1**	**3**	**0.2**	**1**	**29%**	2.10	8.30		52.0	52.0
Lower arm	**1**	**0.5**	**0.5**	**0.04**	**29%**	0.88	8.30	11.28	0.7	52.8
Scoop	**2**	**3**	**0.2**	**0.02**	**29%**	2.10	3.00	5.10	0.5	53.3
Raise arm	**3**	**3**	**0.2**	**0.05**	**29%**	2.10	3.00	5.10	1.3	54.5
Drive to goal	**4**	**2**	**0.2**	**0.5**	**29%**	1.40	3.00	4.40	11.0	65.6
Dump pucks	**5**	**0.1**	**0.5**	**0.05**	**29%**	0.18	3.00	3.18	0.3	65.9



Newton: *Free Body Diagrams & Superposition*

The *structural loop* (see page 3-24) traces the path forces and moments take through a machine's structure and components; accordingly, it also serves as a sort of map to guide the analysis of each element. This allows for a complex structure to be broken down into elements simple enough to be analyzed. The relation between forces and moments transmitted across the boundaries between these individual elements is depicted graphically using a *free-body diagram*. Newton says that in order for a body to remain at rest, the net effect of all forces and moments on the body must be zero. Hence if a machine is drawn in an exploded view, where none of its principle elements are touching, one can represent the flow of forces and moments through the system by drawing force and moment vectors at each elements' interface points. Newton also says that for every action there is an equal and opposite reaction, so the magnitude and direction of the force and moment vectors at any interface between two elements must be equal and opposite.

Consider the simple example of a ladder, which is shown in its assembled state and as an exploded assembly with all the forces shown. There are no moments between members because all the joints are pin joints which are free to rotate. Also, friction forces between the ground and the feet are ignored (maybe its icy!). To determine the magnitude of the five unknown forces A_x, C_x, C_y, N_{yE}, and yD_2, we write the force and moment balance equations for the two legs.

$$\sum F_x = 0 = -C_x + A_x \qquad \sum F_x = 0 = C_x - A_x$$
$$\sum F_y = 0 = -F + C_y + N_{yE} \qquad \sum F_y = 0 = -C_y + N_{yD}$$
$$\sum M_C = 0 = aA_x - cN_{yE} \qquad \sum M_C = 0 = -aA_x + dN_{yD}$$

We could blindly forge ahead with solving these 5 simultaneous equations, but it is easier, apply Occam's razor, to do the force and moment balance for the assembly to easily determine the forces N_{yE}, and N_{yD} and then the forces A_x and C_x:

$$\sum F_y = 0 = -F + N_{yE} + N_{yD} \qquad \sum M_C = 0 = cN_{yE} - dN_{yD}$$
$$N_{yD} = \frac{cF}{d+c} \qquad N_{yE} = \frac{dF}{d+c} \qquad A_x = C_x = \frac{cdF}{a(d+c)}$$

The results are easily incorporated into a spreadsheet, and can be used to design the cross section of the ladder legs, remembering of course that there are two sides to each leg, to resist the applied loads. And of course an appropriate safety factor would have to be applied, for the case where the homeowner who weighs 1000 N steps on the top step carrying 200 N of roof repair materials.

Similarly, complex loads can be broken up into components and their effects analyzed one at a time, which is the foundation for the principle of superposition. Analytically, superposition can be defined in the following manner[1]: *"With certain exceptions (large or plastic deformations), the effect (stress, strain, or deflection) produced on an elastic system by any final state of loading is the same whether the forces that constitute the loading are applied simultaneously or in any given sequence and is the result of the effects that the several forces would produce if each acted singly"*. Philosophically, superposition says that one should never fear, for even the most complex problems can be broken up into little manageable ones.

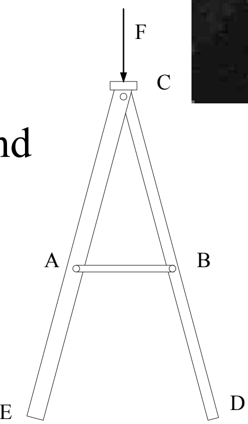
Consider the case of the stepped shaft that is used to isolate radial error motions at one end of a shaft to minimize radial error motions at the other end of the shaft. The stepped shaft is broken into two parts. Starting at the end, the input force is applied, but in order to balance this force, at the other end, an opposite force must be applied. However, these two opposed forces create a couple which would cause the segment to pitch; hence a moment must exist to balance the segment. The second segment's end is where the reaction forces are applied, so a force and a moment with opposite direction are applied. The springs also apply forces. The free end of the segment has no forces or moments, but we will be wanting to determine the deflection at this point. The problem can now be analyzed in two simple parts.

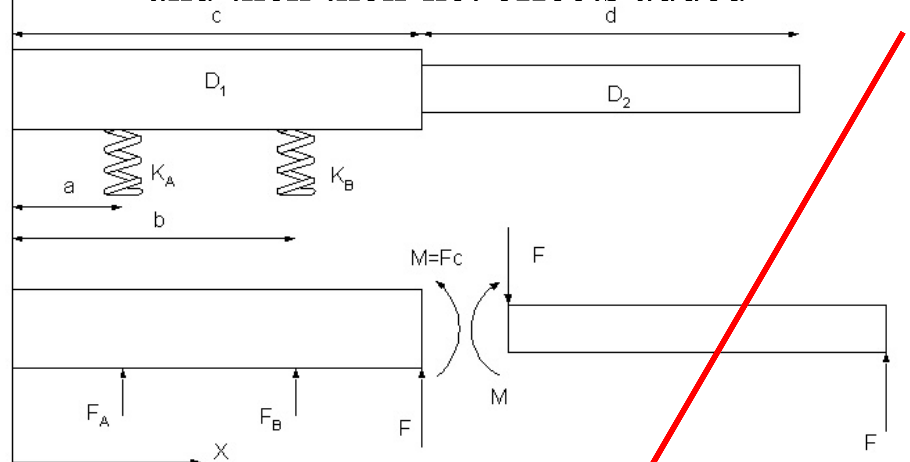
Draw *free body diagrams* for your ***concepts***. How can your **concepts** be divided up into ***modules*** and how are forces on the machine distributed amongst the ***modules***? How are the forces distributed within the ***modules***? Can you use your insight into the distribution of ***modules*** to help you create ideas for the detail design of the ***modules***?

1. R. Roark, W. Young, Formulas for Stress and Strain, 1975, McGraw Hill, New York

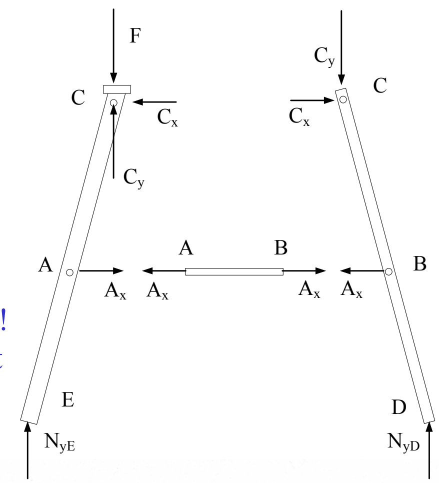
Newton: *Free Body Diagrams & Superposition*

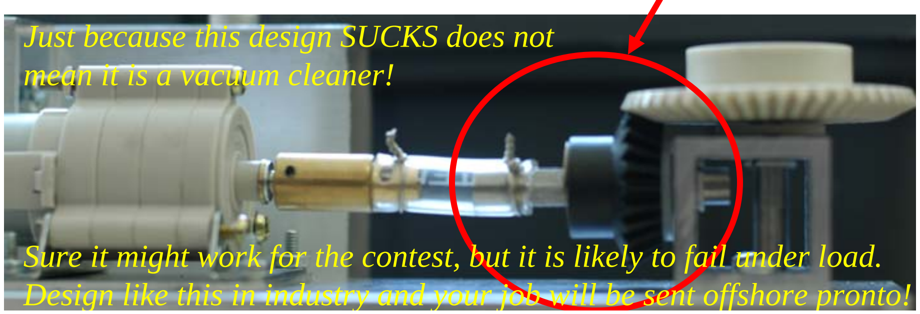
- *Free body diagrams* are a graphical representation of Newton's third law
  - They allow a designer to show ***components*** and their relationship to each other with respect to forces transmitted between them
    - Invaluable for properly visualizing loads on ***components***
    - In order to properly constrain a ***component***, one has to understand how it is loaded and constrained
- *Superposition* allows a complex load to be broken up into ***components*** each of which can be applied one at a time, and then their net effects added

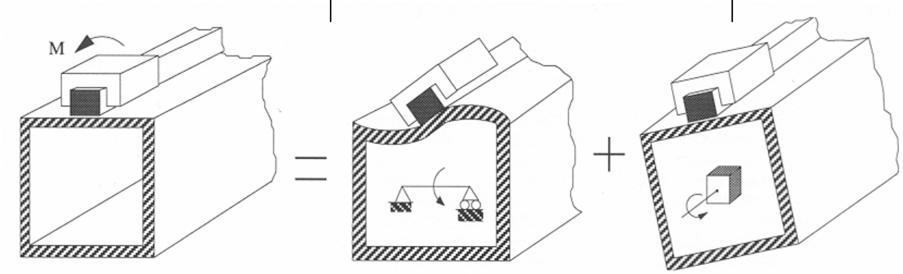




What supports the other end of the shaft to which the gear is attached? How will the gear-tooth radial forces be resisted? A simple FBD of every component can be a critical design synthesis catalyst. FDBs are critical to helping identify how to properly support components! (in a few pages, Saint-Venant will…)







Conservation of Energy...

Einstein showed that Newton's "Laws" were actually "Newton's Pretty Good Estimates" which work well when things are not moving too fast... However, it holds true that *everything that goes into something must be accounted for in one way or another!* Such is the ***Conservation of Energy***.

The units of *energy* are the *Joule (J)*. There are many different physical phenomenon that are all represented by the same units of energy: e.g., thermal, chemical, mechanical, electrical, and fluidic. A mechanical form of energy is the product of *force* (Newtons) and *distance* (meters): For example, 9 Newtons of force (about 2 pounds of force) applied over a distance of 1 meter (about 40 inches) represents 9 N-m of energy (about 80 lb-in). Think about it physically, if you push hard over a long distance, you need use a lot of energy. Another form of mechanical energy is the product of *torque* (N-m) and *rotation*[1] (radians): 10 N-m X $2\pi = 20\pi$ N-m = 20π J.

Energy and matter are equivalent as Einstein showed:

$$e_{\text{energy (Joules)}} = m_{\text{mass (kilograms)}} \times c^2_{\text{speed of light (meters/second)}}$$

All the energy and/or mass that goes into a system, must come out as either energy or mass. In the case of nuclear fission, a neutron goes flying into a uranium atom, the atom splits, little pieces go flying everywhere and heat is also generated. However, if you add up all the mass before and after, after there is a little less mass because some of the mass was converted into energy (heat). This direct conversion of mass into energy only occurs in nuclear reactions. In chemical reactions, energy is liberated from chemical bonds which were storing the energy. Thus when you burn wood and get heat, the mass of the materials going into the reaction, oxygen and wood, equals the mass of the materials afterwards (smoke and ash); The state of the mass changed, even though its quantity did not, and energy (heat) was liberated in the process.

Energy can simultaneously exist in multiple forms: When your body sends sugar into its muscles with a signal to move, out comes force and motion accompanied by heat and waste by- products. The heat represents the efficiency of the conservation of energy, and thus the conservation of energy is often represented as:

$$E_{\text{energy in}} = n_{\text{efficiency}} E_{\text{energy out}}$$

Some people continue to try and "invent" perpetual energy machines, where they claim to get more out than they put in. Invariably, they somewhere forget to account for thermal losses: The efficiency n is always less than 1. For engineering estimate purposes, when the system is slow moving and nearly frictionless, it can be acceptable to assume that η is 1. The *Second Law of Thermodynamics* is always true: *You cannot generate more energy than you put in, all you can do is convert it from one form to another.* This is the essence of all science and engineering. Many thinking people rightly so spend a lifetime truly understanding this very simple principle because it can be applied in so many ways to so many things.

The key to using the conservation of energy is to identify how the energy comes into the system and in what amount, and how it leaves the system. For example, a little force applied to the long end of a simple lever causes it to move a lot, while at the short end of the lever, a larger force is created that moves over a shorter distance. The same applies for solving for the torques and rotations of gears and screws.

When applying the conservation of energy, you will have a force applied over a distance going in that equals a force applied over a distance going out. You can select either force and find the other if you understand the geometric relation created by the constraints of the mechanical system. This is analysis of *geometric constraint*. It's that simple! And the same philosophy holds true for other types of energy calculations.

Look closely to the examples shown which can be powerful tools/catalysts in early feasibility study (Analysis column of FRDPARRC) and design layout stage for your machine! The key to applying the principle of *conservation of energy* is to identify the forces that act on the system and the distance over which they act. Use simple figures to help you identify geometric relations. Use *Free Body Diagrams* to help you account for all the forces and moments in a system and between components.

---

1. Remember, always use radians for angular measure in energy calculations! $2\pi = 360^o$. The units of radians are implied when using π as a unit of angular motion measure.

Conservation of Energy

- What goes in must come out:

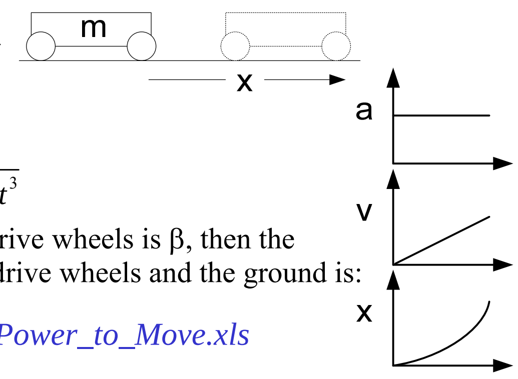
$$n_{\text{efficiency}} \times E_{\text{energy in}} = E_{\text{energy out}}$$

$$E = F_{\text{force out (N)}} \times d_{\text{distance out (m)}}$$

$$E = \Gamma_{\text{torque (or moment) (N-m)}} \times \alpha_{\text{distance (radians)}}$$

For a machine of mass m to move a distance x under constant acceleration in time t:

$$x = \frac{at^2}{2} \qquad v = at \qquad F = ma \qquad P = \frac{Fv}{\eta_{efficiency}}$$



Solving for the power consumed

$$a = \frac{2x}{t^2} \qquad v = \frac{2x}{t} \qquad F = \frac{2xm}{t^2} \qquad P = \frac{4mx^2}{\eta_{efficiency} t^3}$$

If the percent weight of the vehicle over the drive wheels is β, then the minimum coefficient of friction between the drive wheels and the ground is:

$$\mu_{\text{minimum}} = \frac{F}{mg\beta}$$

See *Power_to_Move.xls*

Assume F_{in} and d_{in} are known inputs; find F_{out} and d_{out}:

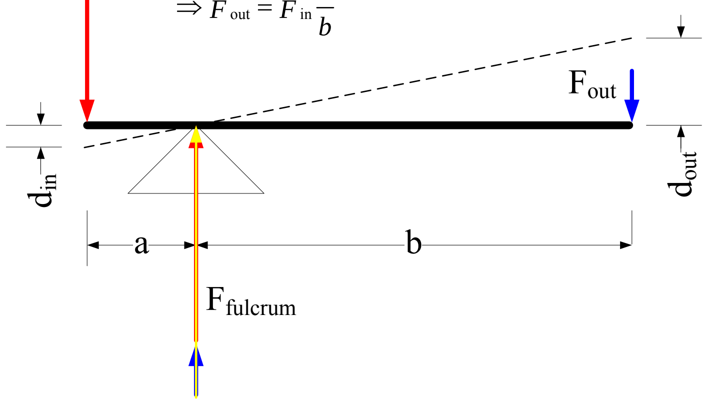
$F_{in} \times d_{in} = F_{out} \times d_{out} \Rightarrow$ 1 equation and 2 unknowns

$\Rightarrow \frac{d_{in}}{a} = \frac{d_{out}}{b} \Rightarrow$ geometric compatability (small angles)

F_{in}

$\Rightarrow d_{out} = d_{in}\frac{a}{b} \Rightarrow$ use in the first equation

$\Rightarrow F_{out} = F_{in}\frac{a}{b}$



Force and Moment Equilibrium (Newton's First Law):

Assume F_{in} is a known input; find F_{out} and $F_{fulcrum}$

$\sum F_y = 0 \Rightarrow -F_{in} - F_{out} + F_{fulcrum} = 0 \Rightarrow$ 1 equation and 2 unknowns

$\sum M_A = 0 \Rightarrow F_{in} \times a = F_{out} \times b \Rightarrow$ 1 equation and 1 unknown

$\Rightarrow F_{out} = F_{in}\frac{a}{b} \Rightarrow$ use in the first equation

$$\Rightarrow F_{fulcrum} = F_{in} + F_{out} = F_{in}\left(1 + \frac{a}{b}\right)$$

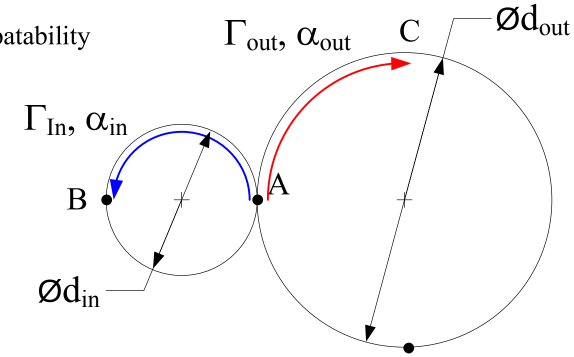
Assume Γ_{in} and α_{in} are known inputs ($\alpha_{in}=\pi$); find Γ_{out} and α_{out}:

$\Gamma_{in} \times \alpha_{in} = \Gamma_{out} \times \alpha_{out} \Rightarrow$ 1 equation, 2 unknowns

$\Rightarrow \pi\frac{d_{in}}{2} = \alpha_{out}\frac{d_{out}}{2} \Rightarrow$ geometric compatability

$\Rightarrow \Gamma_{in} \times \pi = \Gamma_{out} \times \pi\frac{d_{in}}{d_{out}}$

$\Rightarrow \Gamma_{out} = \Gamma_{in}\frac{d_{out}}{d_{in}}$



See *Spurgears.xls*

Assume torque applied to screw is Γ_{in} over one revolution ($\alpha_{in}=2\pi$)

Lead ℓ is defined as distance nut travels in one screw revolution

$\Gamma_{in} \times 2\pi \times \eta_{efficiency} = F_{out} \times \ell \Rightarrow$ 1 equation, 1 unknown

$$\Rightarrow F_{out} = \frac{2\pi\eta\Gamma_{in}}{\ell}$$

See *Screwforce.xls*

Saint-Venant's Principle

In the 19th century applied mathematicians were using the relatively new tool of calculus catalyzed by needs of the industrial revolution to develop theories on the elastic behavior of solids. One problem that often precluded finding rigorous solutions to practical problems was complexity, caused by the desire to include every possible facet of a system in its model. For example, when determining the deflection of a cantilever beam, modelling the local deformations of the force applied to the tip of the beam can make the problem intractable with the added complexity. It would take a gifted applied mathematician with a practical bent to address this issue.

Barré de Saint-Venant was born in 1797 in the castle de Fortoiseau (Seine-et-Marbe). Although he was very gifted at mathematics, the first 26 years of his life led him through many non-academic experiences that certainly honed his resolve for hard work and focus[1]. When he was admitted to university in 1823 without examination, he was ostracized by other students and this apparently further increased his focus and resolve. Upon graduation, he worked as an engineer, and because the infant theory of elasticity had few rigorous solutions for practical engineering application, he became a strong proponent of improving knowledge by coordinating experimental and practical work with theoretical study. Saint-Venant would go on to make many fundamental contributions to the theory of elasticity, but his great philosophical contribution to engineering is his approach to modeling practical problems.

In order to make analytical models more tractable, what is referred to as *Saint-Venant's principle* is often applied, which states that several characteristic dimensions away from an effect, the effect is essentially dissipated. Saint-Venant demonstrated this by using a pair of pliers to squeeze a rubber bar, and the deformation and stress effects become very small 3-5 bar thicknesses away. Saint-Venant's principle is indeed very general and one must be careful when applying it to use it only as a guideline during the design concept and layout phase. Philosophically, we can apply Maxwell's reciprocity to Saint-Venant's principal and observe that if an effect is to dominate a system, it must be applied over 3-5 characteristic dimensions of the system. More rigorous analysis including a sensitivity study of the design parameters should be completed before a design is committed to production.

1. S. P. Timoshenko History of Strength of Materials, Dover Publications, New York, 1983

The applications of Saint-Venant's principle are very apparent in many aspects of machine design from bearings to structures to bolted joints. When mounting bearings to support a shaft, for example, the bearings should be spaced 3-5 shaft diameters apart if the bearings are to effectively resist moments applied to the shaft. In a machine tool structure, for example, if one is to minimize bending, the length of the structure should be no more than 3-5 times the depth of the beam. When bolting components together, in order to make the bolted joint act almost as if it were welded together, the bolts' strain (stress) cones should overlap. The strain cone emanates from 45 to 60 degrees under the bolt head as shown. The strain cones typically overlap if the bolts are spaced less than 3-5 bolt diameters apart.

Consider the problem of a beam of length *L* supported at two points *A* and *B*, for example, bearings, which are distances *a* and *b* respectively from one end. A load *F*, such as caused by coupling errors to another shaft is applied at the other end of the beam. We want to determine the displacement δ of the beam at its origin, such as where a gear may be attached, in order to properly set the center distance between gears. This is a typical model for a shaft, where one wants to determine if radial forces cause too large deflections which might, for example, cause gear teeth to not mesh properly. However, this model is not easily found in most handbooks, so it is good to be able to derive the equations of bending deformation. The loading function *q*, shear *V*, moment *M*, slope α and deflection δ are determined using singularity functions:

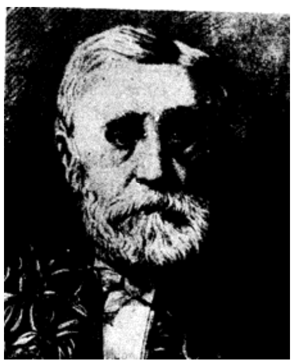
$$\delta(x) = \int \alpha(x)dx = \frac{1}{EI}\left(\frac{F_A\langle x-a\rangle^3 + F_B\langle x-b\rangle^3 + F\langle x-L\rangle^3}{6}\right) + c_1x + c_2x$$

With this equation, one can assume a shaft misalignment of Δ and then determine the force at $x = L$ this creates, and then the resulting displacement at $x = 0$. This allows the design parameters (shaft diameter, length and support locations) to be optimized such that the displacement δ is less than the nominal center distance variation that would be allowed for the gear (or other machine element) that is to be mounted to the shaft.

Complete the details of the analysis and study the results using a spreadsheet and "what-if" scenarios. Keep this spreadsheet handy for your designs!

Saint-Venant’s Principle

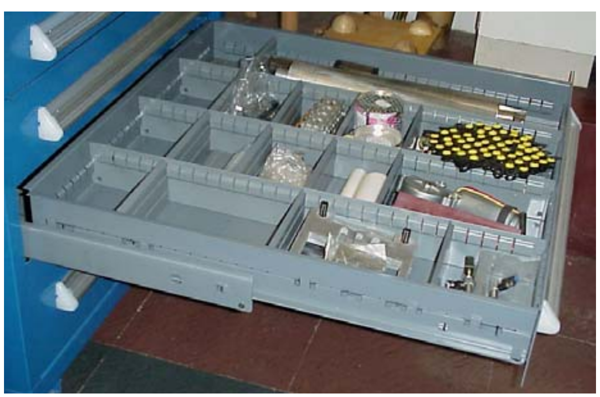
- Saint-Venant did research in the theory of elasticity, and often he relied on the assumption that local effects of loading do not affect global strains
  - e.g., bending strains at the root of a cantilever are not influenced by the local deformations of a point load applied to the end of a cantilever
- The engineering application of his general observations are profound for the development of conceptual ideas and initial layouts of designs:
  - To NOT be affected by local deformations of a force, be several characteristic dimensions away
    - How many seats away from the sweaty dude do you want to be?
    - Several can be interpreted as 3-5
  - To have control of an object, apply constraints over several characteristic dimensions
    - These are just initial layout guidelines, and designs must be optimized using closed-form or finite element analysis

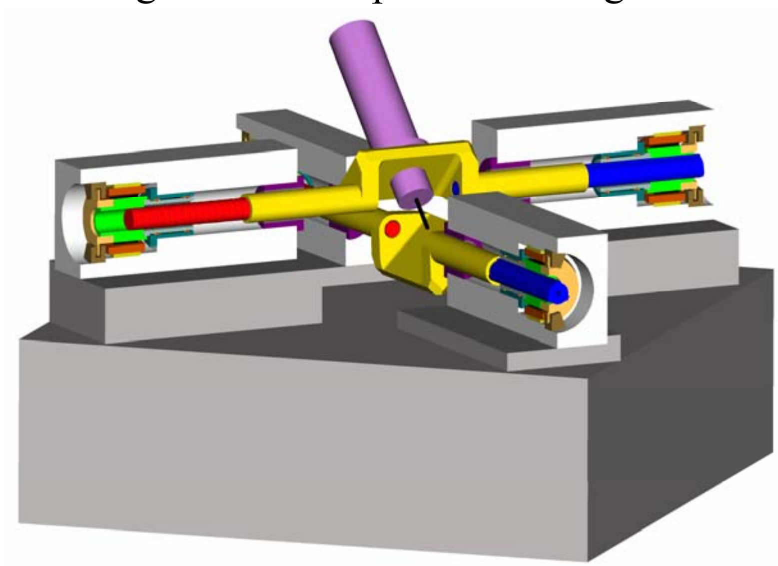


Barré de Saint-Venant

1797-1886

- One of the most powerful principles in your drawer of FUN*da*MENTALS





Saint-Venant's Principle: *Structures*

Saint Venant's principle also applies to structures. If a beam's length is less than 3-5 of its characteristic dimensions, e.g., its height, then bending will be less of an issue and shear deformations become more important. If the beam is long and slender, then shear deformations can be ignored. In a truss, where shear must be transmitted between the top and bottom chords, the diagonal members should be spaced less than the height of the beam. Similar attention to proportions upon initial layout of a design can help the design engineer avoid trouble as the analysis of a design progresses.

For example, one of the principle issues in the design of structures is the influence of structural deformations on other components. Consider a linear motion system, where the carriage has to be stiff so deformations do not cause the bearings to bind. A solid axle is used so the wheels will turn together to help prevent yawing (jambing). Superposition is used to divide the system into the chassis and the axle for the wheels. The bearing clearance must be greater than the product of the bearing length and the difference of the slope of the chassis deflection at the point at which the bearings are attached, and the local slope of the axle. Starting with the chassis structure, $F_A = F_B = wL_1/2$ and using singularity functions[1]:

$$q(x) = F_A\langle x\rangle_{-1} - w\langle x\rangle^0$$

$$V(x) = -\int q(x)dx = -F_A\langle x\rangle^0 + w\langle x\rangle^1 + c \quad V=0@x=0; c=0$$

$$M(x) = -\int V(x)dx = F_A\langle x\rangle^1 - \frac{w\langle x\rangle^2}{2} + c \quad M=0@x=0; c=0$$

$$\alpha(x) = \frac{1}{EI}\int M(x)dx = \frac{1}{EI}\left(\frac{F_A\langle x\rangle^2}{2} - \frac{w\langle x\rangle^3}{6} + c_1\right)$$

$$\delta(x) = \int \alpha(x)dx = \frac{1}{EI}\left(\frac{F_A\langle x\rangle^3}{6} - \frac{w\langle x\rangle^4}{24} + c_1 x + c_2\right)$$

The boundary conditions are $\delta(0) = \delta(L_1) = 0$, therefore $c_2 = 0$ and:

$$c_1 = -wL_1^3/24$$

For the axle, $b = a$, and thus $F_{w1} = F_{w2} = F_A$,:

$$q(x) = F_{w1}\langle x\rangle_{-1} - F_A\langle x-a\rangle_{-1} - F_B\langle x-b\rangle_{-1}$$

$$V(x) = -F_{w1}\langle x\rangle^0 + F_A\langle x-a\rangle^0 + F_B\langle x-b\rangle^0$$

$$M(x) = F_{w1}x - F_A\langle x-a\rangle^1 - F_B\langle x-b\rangle^1$$

$$\alpha(x) = \frac{1}{EI}\left(\frac{F_{w1}x^2}{2} + \frac{F_A\langle x-a\rangle^2}{2} + \frac{F_B\langle x-b\rangle^2}{2} + c_1\right)$$

$$\delta(x) = \frac{1}{EI}\left(\frac{F_{w1}x^3}{6} + \frac{F_A\langle x-a\rangle^3}{6} + \frac{F_B\langle x-b\rangle^3}{6} + c_1 x + c_2\right)$$

The boundary conditions are $\delta(0) = \delta(L_2) = 0$, therefore $c_2 = 0$ and:

$$c_2 = -\frac{wL_1}{12L_2}\left(L_2^3 - (L_2-a)^3 - (L_2-b)^3\right)$$

The formulas are implemented in the spreadsheet *Axledefl.xls*, which can be easily modified for other similar problems. Is this a good design? Would the design be better if two axles were used where each was supported by its own bearings? What kind of analysis is needed to ensure that the deformations in the system do not overload the bearings? Imagine the care that must go into a weight sensitive machine like an airplane or satellite!

How might this model apply to your ***concepts***? Highlight bearings and the structures that affect them and consider how you can analyze them.

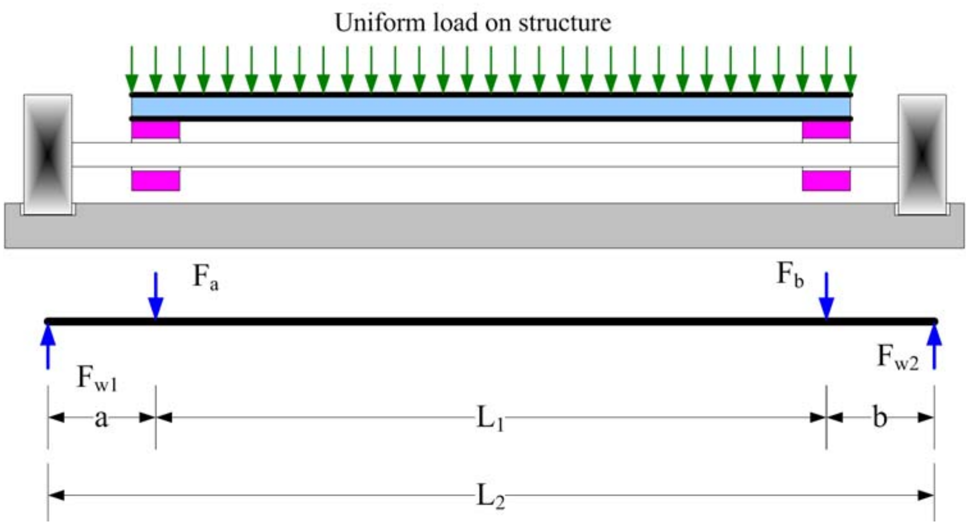
---

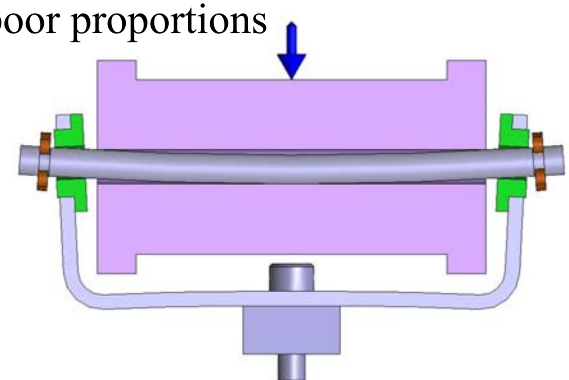
1. Loads are represented by the loading function q(x) and are "turned on" at a position by multiplying them by the position at which they occur using a brackets term. Distributed loads (N/m) are multiplied by $\langle\rangle^0$, forces (N) by $\langle\rangle_{-1}$, and moments (N-m) by $\langle\rangle_{-2}$ (see S. Crandahl, N. Dahl, T. Lardner An Introduction to the Mechanics of Solids, pp 164-172, McGraw-Hill, Inc. NY, NY). The terms $\langle\rangle$ are integrated according to:

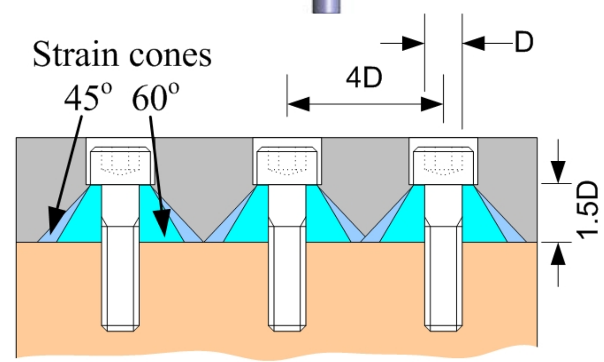
$$\int_{-\infty}^{x}\langle x-a\rangle_{-n} = \langle x-a\rangle_{-n+1} \qquad \int_{-\infty}^{x}\langle x-a\rangle^{n} = \frac{\langle x-a\rangle^{n+1}}{n+1}$$

Saint-Venant's Principle: *Structures*

- To NOT feel something's effects, be several characteristic dimensions away!
  - If a plate is 5 mm thick and a bolt passes through it, you should be 3 plate thicknesses away from the bolt force to not cause any warping of the plate!
    - Many bearing systems fail because bolts are too close to the bearings
  - Beware the strain cone under a bolt that deforms due to bolt pressure!
    - Strain cones should overlap in the vicinity bearings to prevent wavy deformations
    - BUT check the design's functional requirements, and only use as many bolts as are needed!
- To DOMINATE and CONTROL something, control several characteristic dimensions
  - If a column is to be cantilevered, the anchor region should be 3 times the column base area
    - Too compliant machines (lawn furniture syndrome) often have poor proportions
    - Diagonal braces can be most effective at stiffening a structure







Saint-Venant's Principle: *Bearings*

The initial layout of a machine's ***concept*** must include the proposed points of contact and motion between components. The relative distance between bearings, the location of applied loads, and the sizes of the structural components are of critical importance. In any bearing application, an analysis of the loads must be done, but too often designers do the analysis after they have already spent hours on the CAD station creating a lot of detail. Being able to create a first order solid model of the system with appropriate proportions is an invaluable skill. Designers should continually learn about proportions by looking at real world devices from bikes to construction equipment.

Bearings are often the principle point of failure in a machine because relative motion between their elements causes wear which is often more complex and difficult to control than structural fatigue. Bearings' detailed design considerations are discussed in Chapters 10 and 11. With respect to *Saint Venant's Principle*, failure in bearing systems can occur due to direct overloading, or indirect overloading. The former occurs when the bearing is simply undersized for the directly applied load. The latter occurs most often when the designer did not consider the effect of the rest of the system on the bearings.

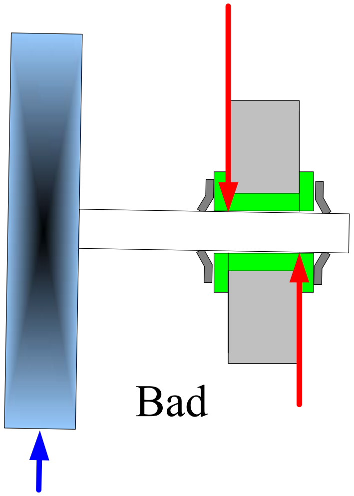
One of the most common bearing systems that we can all relate to are the bearings in a drawer. When the bearings are spaced at the sides of the drawer, their engagement length decreases as the drawer is pulled out. As the ratio of their separation to their engagement length decreases, there comes a point where they may jam, and you curse the cheap design. Better quality drawers have a central guide system that prevents this problem, but this does not work well for large heavily loaded drawers. Telescoping ball bearing drawer guide systems ensure that an acceptable engagement length is maintained through the use of multiple stages. In addition ball bearings reduce the tendency of friction to create a moment instability that leads to jamming.

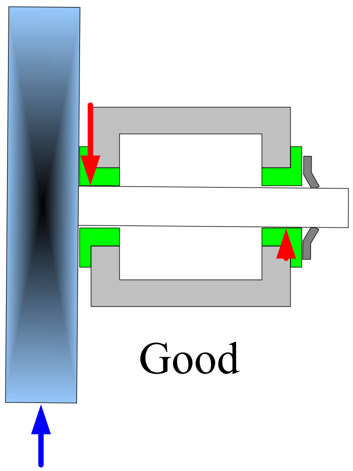
Consider the mounting of a wheel to a shaft that is to be held by bearings. The critical design parameters are the distance from the wheel to the first bearing, the distance between the first and second bearings, the diameter of the shaft, *and* the support structure and its loading. In the first case, if the wheel is cantilevered out far from the first bearing, it can cause the bearing to be overloaded. The cantilever distance should be less than 2 shaft diameters in order to prevent angular deflections of the shaft from taking up all that is allowed for the front bearing which acts as a fulcrum. Next, the shaft in the region between the first and second bearings acts as a lever. If the lever is made very long to minimize forces on the bearings, then there is the risk that it will deform too much and cause the bearings to overload. In order for the second bearing to have dominance over the system, but not allow the shaft to bend too much, it should be at least 3-5 shaft diameters from the first bearing. If a larger spacing is desired, perhaps to reduce the number of ***components***, then one must consider the effects of deformations of the structure surrounding the bearings.

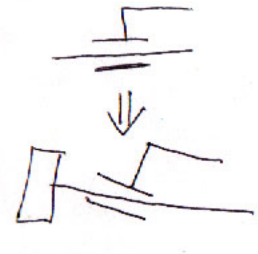
Next consider the cross section through a vehicle where the shaft, wheels, and bearings appear to be sized and located with reasonable dimensions. Analysis might show that the stresses and deflections in the shaft and bearings are reasonable; however, if the chassis of the vehicle is very compliant, a heavy load applied in the middle might cause bending and a large angular deflection at the bearings. This would cause the shaft to pivot in the bearings, take up all the radial clearance, and cause bearing failure. The designer can either increase the stiffness of the vehicle chassis, or use two axles and two sets of bearings. In the latter case, complexity is increased, but large deflections can be accommodated. In steel and paper mills, the shafts are so long and heavily loaded, that the only way angular deflections can be accommodated is with the use of spherical bearings. When first introduced, spherical bearings were a disruptive technology; pursuing the details to the point of understanding the fundamentals can lead to important new inventions!

Look at the proportions of your ***modules*** and ***components*** and where loads are applied and decide where to put the bearings so they can have dominance over applied loads. Can deformations in the structure cause the bearings to be overloaded? If the loads are applied at a distance far away to the bearings in proportion to their spacing, do not panic, but take note that before you proceed further with the design, you will have to do detailed calculations to ensure the bearings can handle the loads. Also consider how you will attach the bearings to the structure and take note of local deformation effects such as strain cones under bolt heads.

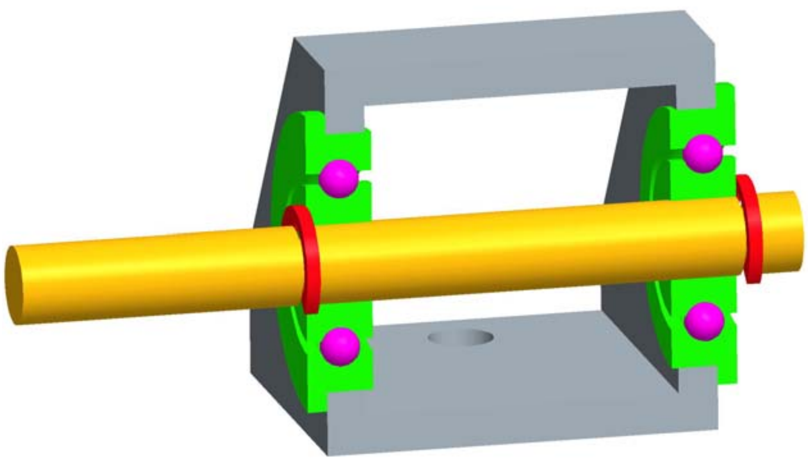
Saint-Venant's Principle: *Bearings*

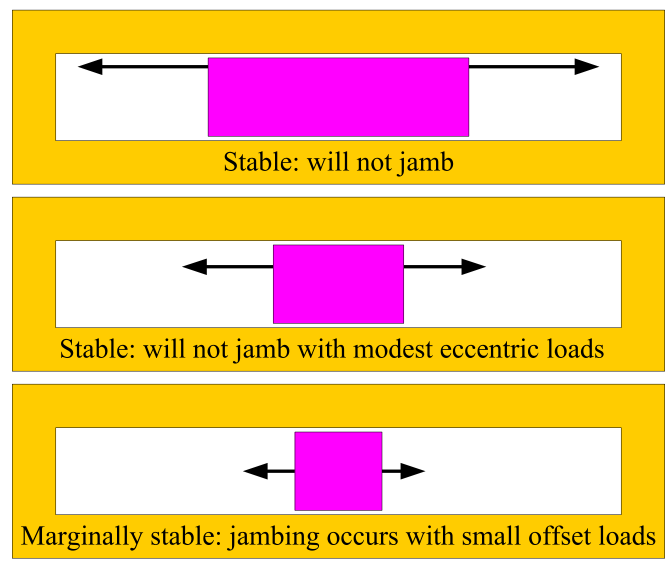






- Saint-Venant: *Linear Bearings:*
  - Make friction (μ) low and L/D > 1, 1.6:1 very good, 3:1 awesom
  - Every year some students try L/D < 1 and their machines jam!
    - Wide drawers guided at the outside edges can jamb
    - Wide drawers guided by a central runner do not!
    - If L/D < 1, actuate both sides of the slide!
- Saint-Venant: *Rotary Bearings*:
  - L/D > 3 if the bearings are to act to constrain the shaft like a cantilever
  - IF L/D < 3, BE careful that slope from shaft bending does edge-load the bearings and cause premature failure
  - For sliding contact bearings, angular deformations can cause a shaft to make edge contact at both ends of a bearing
    - This can cause the bearing to twist, seize, and fail
    - Some shaft-to bearing bore clearance must always exist





Saint-Venant's Principle: *Bearings*

Let us now consider a more likely design model of the shaft supported by bearings, where the support points A and B are springs of constant K_A and K_B and the shaft diameter D_1 changes to D_2 a distance c from the origin (where $c>b$). Superposition is used to break the problem up into that of a cantilever beam D_2 and length d that is attached to a beam D_1 that is supported by springs. The deflection of the beam D_2 at its tip equals the deflection of a cantilever beam with a force F applied at its end plus the deflection of the end of beam D_1 subject to a force F and a moment Fd and the product of the slope $\alpha(c)$ of the end of beam D_1 and the length d. Page 3-4 described the process of creating a *free-body-diagram* for the system to show forces and moments on each section. The first step is to find the forces at the support points which is done with a simple force and moment balance:

$$F_A = F\left(\frac{c+d-b}{b-a}\right) \qquad F_B = -F\left(\frac{c+d-a}{b-a}\right)$$

The loading function and integration steps for the D_1 beam are:

$$q(x) = F_A\langle x-a\rangle_{-1} + F_B\langle x-b\rangle_{-1} + Fd\langle x-c\rangle_{-2} + F\langle x-c\rangle_{-1}$$

$$V(x) = -F_A\langle x-a\rangle^0 - F_B\langle x-b\rangle^0 - Fd\langle x-c\rangle_{-1} - F\langle x-c\rangle^0$$

$$M(x) = F_A\langle x-a\rangle^1 + F_B\langle x-b\rangle^1 + Fd\langle x-c\rangle^0 + F\langle x-c\rangle^1$$

$$\alpha(x) = \frac{1}{EI}\left(\frac{F_A\langle x-a\rangle^2}{2} + \frac{F_B\langle x-b\rangle^2}{2} + Fd\langle x-c\rangle^1 + \frac{F\langle x-c\rangle^2}{2} + c_1\right)$$

$$\delta(x) = \frac{1}{EI}\left(\frac{F_A\langle x-a\rangle^3}{6} + \frac{F_B\langle x-b\rangle^3}{6} + \frac{Fd\langle x-c\rangle^2}{2} + \frac{F\langle x-c\rangle^3}{6} + c_1x + c_2\right)$$

The boundary conditions are $\delta(a) = -F_A/K_A$ and $\delta(b) = -F_B/K_B$, hence the constants C_1 and C_2 are:

$$c_1 = \frac{EI}{b-a}\left(\frac{-F_B}{K_B} + \frac{F_A}{K_A}\right) - \frac{F_A(b-a)^2}{6} \qquad c_2 = \frac{-F_A EI}{K_A} - c_1 a$$

Using *Bearings_rotary_spacing.xls* to examine "what if" scenarios, you will find that for shaft diameters D_1 and D_2, with $a = D_1$, $b = 4D_1$, $c = 5D_1$, $d = 2D_1$ that the ratio of the deflections at the left end/right end becomes very small 0.049 (there is less and less gain to making the supports further and further apart); therefore, error motions at the end of the beam D_2 have an insignificant effect on the error motion of the beam at its beginning. This means, for example, that radial error motion of a motor coupled to the end of D_2 will have a minimal effect on the radial error motion of a gear attached to the beginning of shaft D_1. This relatively simple analysis allows the design engineer to effectively use a shaft as its own coupling system to minimize errors. In a precision machine, this analysis would be used to determine if the radial error motions of a ballscrew nut caused significant errors in a linear motion slide.

Next, consider bearing spacing in width and length and the potential for jamming, like in a drawer. A simple force and moment balance yields the following expression for the force required to pull the drawer:

$$\sum F_Y = 0 = F_{B1} - F_{B2} \Rightarrow F_{B1} = F_{B2} = F_B$$

$$\sum F_X = 0 = -\mu F_B - \mu F_B - \mu W + F_{Pull} \Rightarrow -2\mu F_B - \mu W + F_{Pull} = 0$$

$$\sum M_{@cg} = 0 = -\mu F_B a + \mu F_B a + F_B b + F_B b - F_{Pull} c \Rightarrow 2F_B b - F_{Pull} c = 0$$

$$F_{Pull} = \frac{\mu b W}{b - \mu c} = \frac{W}{1/\mu - c/b}$$

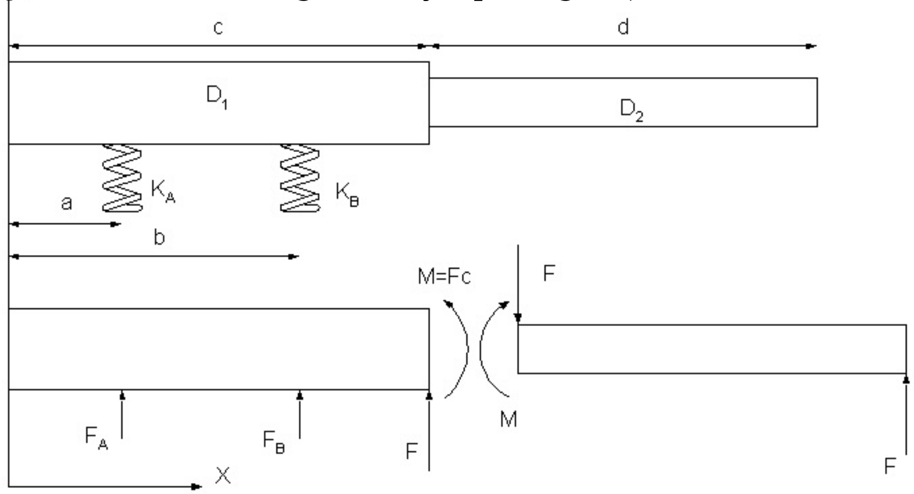
$$F_B = \frac{\mu c W}{2(b-\mu c)} = \frac{W}{2\left(b/\mu c - 1\right)}$$

See *Bearings_linear_spacing.xls*: If a/b is large, and the pull force is offset from the load (weight) on the carriage, excessive loading of the bearings can occur, resulting in early bearing failure. In order to avoid this problem, one can actuate a system from both sides, which is done, for example, in photocopier shuttles.

Take a close look at your ***concepts*** and assess where precision motion is required and where misalignments could cause error motions to be imparted to the system. How can bearing placement and shaft lengths be strategically selected in order to minimize system sensitivity to error motions?

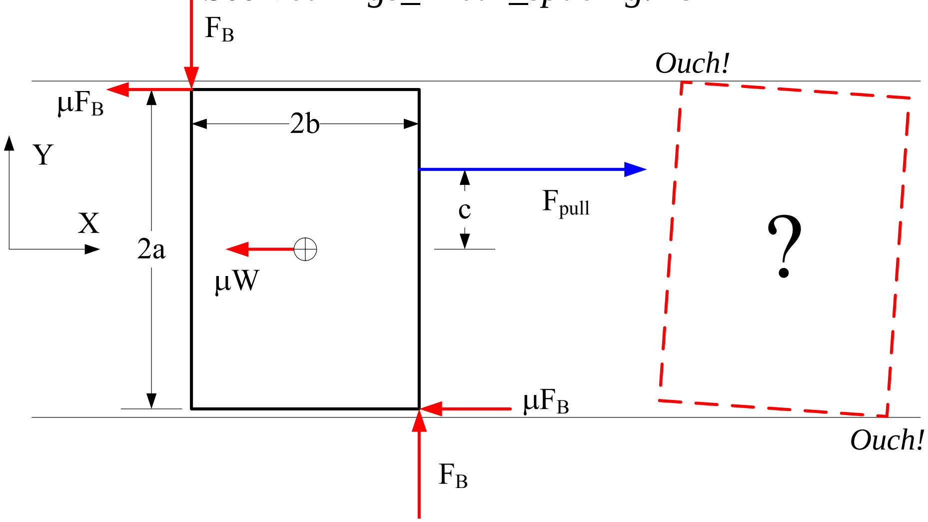
Saint-Venant's Principle: *Bearings*

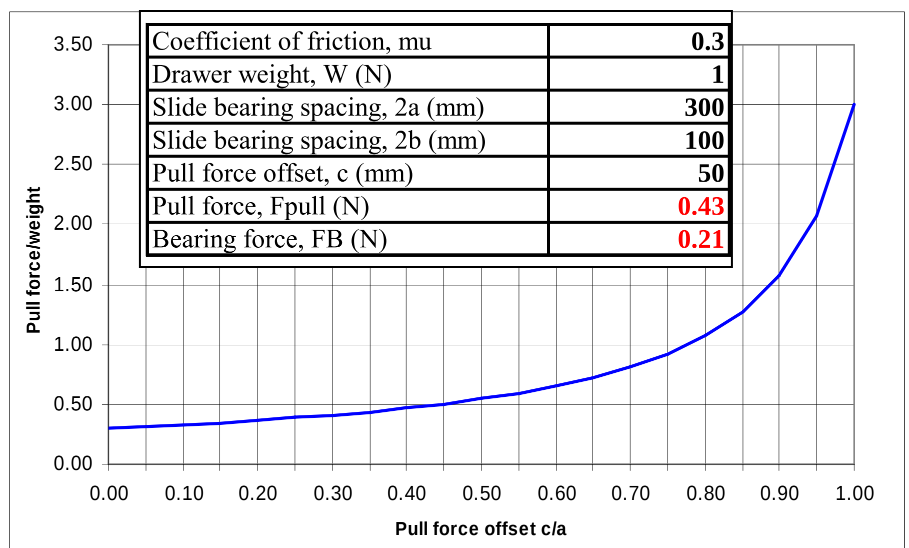
- Model of a shaft supported by bearings: Minimize the deflection of the ends of the beam
  - See *Bearings_rotary_spacing.xls*)



Design Parameters	Values	
a (mm)	50	
b (mm)	20	
c (mm)	250	
d (mm)	100	
Diameter, D_1 (mm)	15	
Diameter, D_2 (mm)	10	
Bearing radial spring constant, KA (N/mm)	2.00E+02	
Bearing radial spring constant, KB (N/mm)	2.00E+02	
Modulus, E (N/mm^2)	6.70E+04	
Tip force, F (N)	10.00	
Moment of inertia, I_1 (mm^4)	2.49E+03	
Moment of inertia, I_2 (mm^4)	4.91E+02	
Spring force, FA (N)	-110.00	
Spring force, FB (N)	100.00	
End deflection of just D_2 segment (mm)	1.01E-01	
End slope of just D_2 segment (rad)	1.52E-03	
Ratio (deflection left end)/(deflection right end)	-0.103	
Position along beam: 0, a, (a+b)/2, b, c, (c+d)	deflection (mm)	slope (rad)
0	-1.20E+00	3.51E-02
50	5.53E-01	3.54E-02
35	2.39E-02	3.52E-02
20	-5.03E-01	3.51E-02
250	7.91E+00	3.78E-02
350	1.17E+01	3.93E-02
Bearing gap closure (for sliding contact bearing supports)		
Bearing width, wb (mm)	5.00	
Diametral clearance loss at first bearing (a) (mm)	0.177	
Diametral clearance loss at first bearing (b) (mm)	0.176	

- Model of the effect of bearing width, friction, and length spacing on the actuation force (drawer jamming)
  - See *Bearings_linear_spacing.xls*





Coefficient of friction, mu	0.3
Drawer weight, W (N)	1
Slide bearing spacing, 2a (mm)	300
Slide bearing spacing, 2b (mm)	100
Pull force offset, c (mm)	50
Pull force, Fpull (N)	0.43
Bearing force, FB (N)	0.21

The Golden Rectangle

The Golden Rectangle was discovered by Pythagoras in ancient Greece to be a rectangle whose sides are in proportion, such that when a square is cut from the rectangle, the remaining rectangle has the same proportions:

$$\frac{a}{b} = \frac{b}{a-b}$$
$$a^2 - ab - b^2 = 0$$
$$a = \frac{1+\sqrt{5}}{2} \approx 1.61803398874989 \quad @b=1$$

If the pattern of subtracting squares continues, and a spiral curve is drawn by linking together circular arcs whose centers are the corners of the squares, the spiral continues to infinity: the chambered nautilus' shell is a Golden Spiral![1] As shown in the Disney's *Donald Duck in Mathmagic Land*[2], there are many other shapes in natures that are framed by the Golden Rectangle and Spiral. Ask yourself what would happen if the second root of the quadratic equation above were taken? a = -0.61803398874989, but b=1 is now the long side, so the ratio of b/a now becomes -1.61803398874989! Maybe a negative *Golden Spiral* is the key to an alternate universe? Maybe this is the form matter takes as it spirals into a *black hole* in our galaxy and then spirals out into a *white bump*? Think of your passion for investigating nature and how its beauty also incites form and function, and that your heart is a golden spiral itself[3]

This sort of helical thinking is a very important part of creative design, for it trains the brain to never refrain and avoid the same; thereby finding new designs you might otherwise miss! Accordingly, the Greeks believed that the Golden Rectangle was a magical feature in nature that was a fundamental building block of perfection. However, as with all the fundamental principles discussed in this Chapter, they represent guidelines to consider, and catalysts for thought. There are very few sacrosanct laws of design.

Fortunately, the Golden Rectangle can help to help lay out a design of a machine. When designing a machine, we can start from the outside, with a sketch of the overall envelop, and work inwards, or we can start from within, with the critical ***module***, and design outwards. The former tends to force the inner workings to be cramped, and sometimes this leads to degradation of performance. The latter tends to result in machines that are too spacious. A compromise from the beginning can be achieved with either method by initially sketching concepts while keeping the proportions of the Golden Rectangle in mind. Saint-Venant's principle also should be considered when initially laying out proportions and spacing of machine elements, whereas the Golden Rectangle helps us with overall proportions of systems.

When you initially sketch a design to layout the principle ***modules***, how can you create a realistic rapid sketch, such that your ***concept*** has a reasonable chance of being able to be engineered? All too often initial quick sketches seem to represent great ideas, but when you get down to doing the detail, you often find that you do not have enough room to make things work. We can thus say that the purpose of sketching designs with the proportions of the Golden Rectangle is to help the design engineer initially sketch ***concepts*** that have a greater chance of being realizable.

Consider the design of a vehicle that may have to climb hills, like those in a robot design contest. The sketch shows a rectangle drawn whose corners are the wheel-to-ground contact points, and its center is the expected center of gravity of the machine. The absolute maximum angle α the vehicle can climb before it tips over is that which causes the center of gravity to become vertically aligned with the rear wheel. As the red line on the graph shows, the angle changes slowly with vehicle aspect ratio, until the proportions of the Golden Rectangle are reached, at which it starts to fall off more rapidly. The proportions of the Golden Rectangle are thus often useful when asked "how much" and the reply is "just enough".

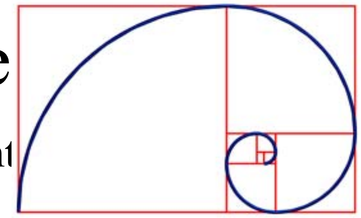
Look at your design concept sketches and evaluate their proportions. Where proportions are very different from the *Golden Rectangle*, do not necessarily rush and change them, but rather use this as a catalyst to more carefully scrutinize them.

---

1. Drawing quarter circles in each square of a *Golden Rectangle* creates a *Fibonacci Spiral*, which is the spiral of shells, the patterns in flowers, pine cones, cauliflower....see http://www.moonstar.com/~nedmay/chromat/fibonaci.htm. Each successive radii in a Fibonacci spiral is that of a *Fibonacci series*: Each number in a Fibonacci series is equal to the sum of the two numbers that came before it.
2. This is a great way for an engineer to get a date, by asking someone to watch it with you, and if they are not interested, they are not worth being with anyway!
3. G.D. Buckberg, MD, "Basic Science Review: The helix and the heart", Jou. Thoracic and Cardiovascular Surgery, Vol. 124, No. 5, Nov. 2002, pp 863-883. See http://jtcs.ctsnetjournals.org/cgi/content/full/124/5/863 for videos and other images.

http://encarta.msn.com/media_461550503_761579463_-1_1/Leonardo_Fibonacci.html

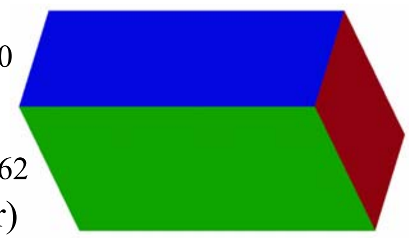
The Golden Rectangle



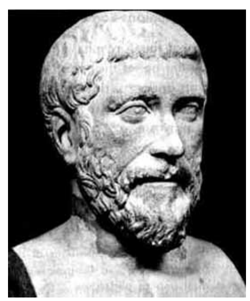
Leonardo Fibonacci (1170?-1240?)

See http://lady-macbeth.mit.edu/2.007/geekdate/

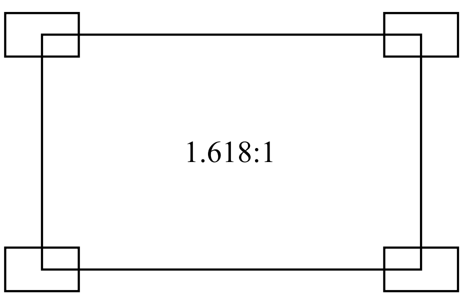
- The proportions of the *Golden Rectangle* are a natural starting point for preliminary sizing of structures and elements
  - *Golden Rectangle:* A rectangle where when a square is cut from the rectangle, the remaining rectangle has the same proportions as the original rectangle: a/1 = 1/(a-1)
    - See and study *Donald in Mathmagic Land!*
  - Try a *Golden Solid:* 1: 1.618: 2.618, & the diagonal has length 2a = 3.236
  - Example: Bearings:
  - The greater the ratio of the longitudinal to latitudinal (length to width) spacing:
    - The smoother the motion will be and the less the chance of walking (yaw error)
    - First try to design the system so the ratio of the longitudinal to latitudinal spacing of bearing elements is about 2:1
    - For the space conscious, the bearing elements can lie on the perimeter of a golden rectangle (ratio about 1.618:1)
    - The minimum length to width ratio should be 1:1
    - To minimize yaw error
    - Depends on friction too

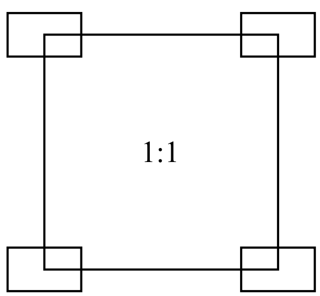


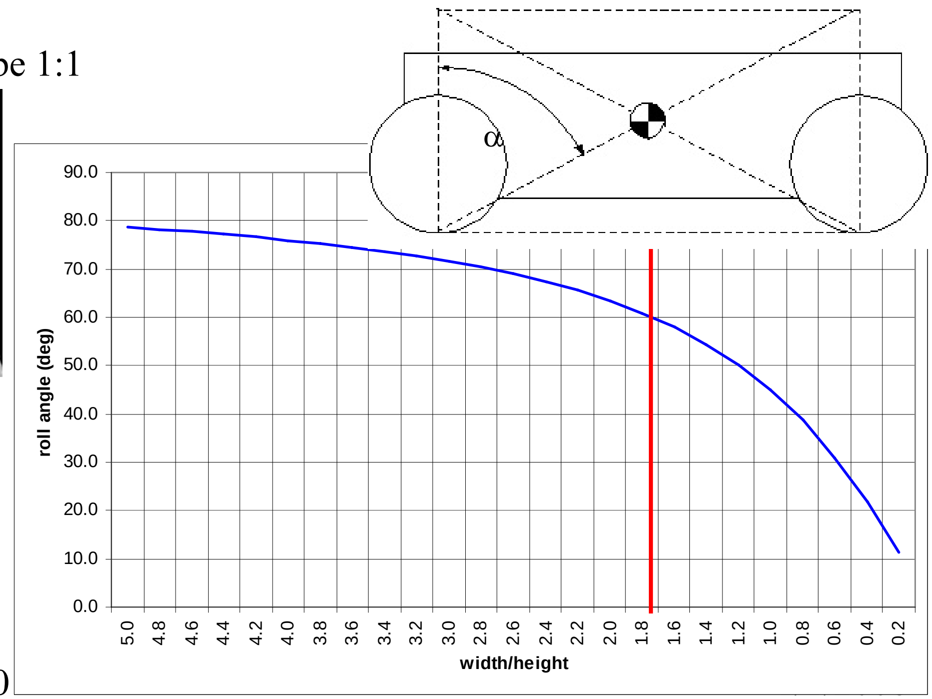
Pythagoras of Samos
569 BC-475 BC



http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Pythagoras.html







Abbe's Principle

Imagine sailing on a boat in rough weather. If you look at the base of the mast you perceive a rolling motion; however, when you look at the top of the mast against a reference point in the sky or on the horizon, it appears to be moving back and forth a lot more. This amplification of angular motion to create large translational motions is one of the foremost principles in the design of precision and robust machines, and yet it was only 150 years ago that the implications of this principle became fully appreciated. It all formally started when Carl Zeiss established himself in the 19th century as a preeminent designer and manufacturer of precision microscopes. When his craftsmanship and intuition reached a limit, he sought the analytical help of Dr. Ernst Abbe. Dr. Abbe developed powerful analytical tools for the design and manufacture of precision instruments, but the most enduring relates to angular errors causing increasing translational errors as one moves further from the source. In other words, the translational error δ in a system at a distance L from an angular error ε (pivot) is $\delta = L\sin\varepsilon$. This is also referred to as a *sine error.* The corollary is the *cosine error*, which as shown implies a far less direct effect, but one that is still important in very high precision systems, such as photolithography.

The *Abbe Offset* is the distance between the axis of measurement, and the axis of the intended motion to be measured. Hence the *Abbe Error* is the product of the Abbe Offset and the sine of the angular error in the system. The source of the angular error is typically geometric error motions in moving mechanical components. The implication of this seemingly innocent observation on the design of instruments and machines are indeed profound:

- Always try to place the measurement system as close to the line of action (the process) as possible.
- Always try to place bearings and actuators as close to the line of action (the process) as possible.

Strictly speaking, angular errors in a bearing's motions that produce appreciable translation errors at the work zone are not Abbe errors. Abbe errors are parallax errors resulting from the measurement system not being colinear with the axis being measured. The former should be called *sine errors*, or errors resulting from the *Abbe effect*[1].

A profound effect of this seemingly trivial type of error caused a huge change of fortunes in the semiconductor equipment manufacturing industry. In the 1980's, GCA Corp. was the world's foremost manufacturer of wafer steppers, or steppers, which are machines used to project circuit patterns onto silicon wafers. GCA was renowned worldwide for making the most accurate machines. Their competitors, Nikon and Cannon made good machines, but could not compete with GCA for making the finest line widths. When it came time to introduce a new generation of machines, however, GCA merely scaled an existing design concept, which used a microscope mounted off axis of the projection lenses to align the wafer to the machine. What used to not be a significant Abbe Offset became significant for the new generation of machines, because the source of the angular error was thermal growth. When the machine was turned on, it was accurate, but as the machine warmed up, an angular error was generated by the heat, and its product with the distance between the alignment microscope's optical axis and the optical axis of the projection lens caused different layers in the wafer to become misaligned. Thus even though the GCA machine had far higher resolution than the Nikon or Canon machines, it drifted with time. Companies would buy a few GCA machines for the development lab to work on next generation devices, but they would buy dozens of Nikon or Canon steppers for production[2]

The philosophy of Abbe's Principle also extends to other types of measurements. For example, when measuring temperature, it is important to place the temperature sensor as close as possible to the process to be measured. The idea is the same for pressure, flow, voltage, current, etc. In each case, the further away you are from the process to be measured, the greater the chance for errors to reduce the accuracy of the measurement.

Review your designs and see where critical motions occur with respect to the controlling mechanism. Can your concept be made more robust by minimizing potential sine errors? Estimate the magnitude of the *sine* and *cosine errors* in your concept.

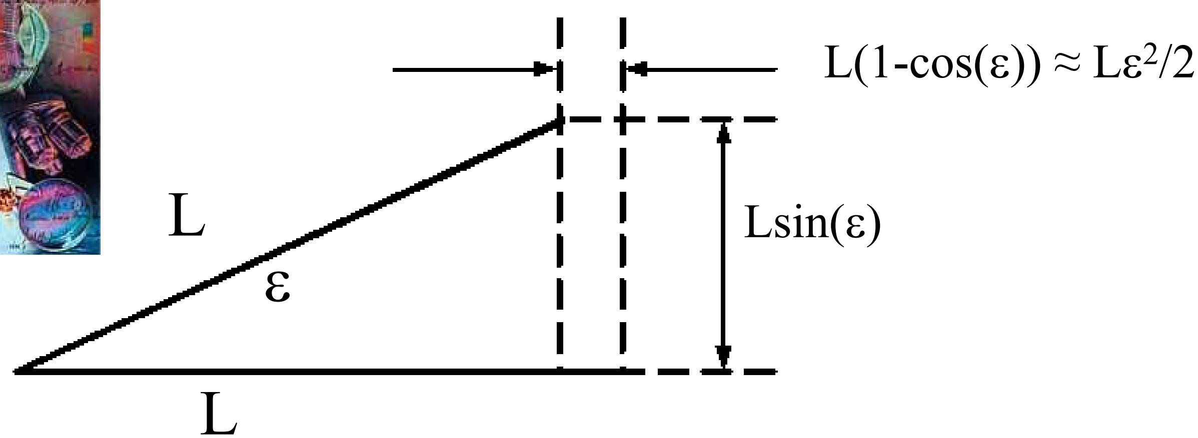
---

1. See for example, J. B. Bryan, "The Abbe Principle Revisited-An Updated Interpretation," Precis. Eng., July, 1989, pp. 129-132.
2. S. Stone, "Flexure Thermal Sensitivity and Wafer Stepper Baseline Drift," paper presented at SPIE OPTCON 1988, Precision Instrument Design Section, Nov., 1988. For a detailed case study discussion, see A. Slocum, Precision Machine Design, pp 541-545, 1996, SME, Dearborn, MI

Abbe's Principle

- In the late 1800s, Dr. Ernst Abbe (1840-1905) and Dr. Carl Zeiss (1816-1888) worked together to create one of the world's foremost precision optics companies: Carl Zeiss, GmbH (http://www.zeiss.com/us/about/history.shtml)
- The Abbe Principle (*Abbe errors*) resulted from observations about measurement errors in the manufacture of microscopes:
  - *If errors in parallax are to be avoided, the measuring system must be placed coaxially with the axis along which the displacement is to be measured on the workpiece*
    - Strictly speaking, the term *Abbe error* only applies to measurement errors
- When an angular error is amplified by a distance, e.g., to create an error in a machine's position, the strict definition of the error is a *sine* or *cosine* error

From www.zeiss.com



Abbe's Principle: *Locating Components*

Consider the task of measuring a shaft, where one can use a micrometer or a dial caliper. The dial caliper is quick and easy to use because with the simple action of your thumb, you can slide the caliper open so it can be placed around the object, and then you can slide it closed to make contact with the object and make the measurement. As can be felt with a cheap caliper, any rocking motion in the caliper head will cause a measurement error. In addition, forces on the caliper jaws cause deformations in the structure and the bearings that support the head. Fortunately, the long range of travel of the caliper head makes it a very versatile measuring instrument.

The micrometer, on the other hand, requires a much more laborious turning of the screw to adjust it to fit over the part, and you have to carefully position it to ensure that you are contacting the part properly. The micrometer, however, is probably 10x more accurate than the dial caliper for the simple reason that the mechanism of motion and measuring are completely in-line with the measurement to be made. In this case, the Abbe offset is zero. However, the range of measurement motion is very limited, making the micrometer a specialized instrument.

Thus we can see the dilemma that design engineers often face: to make something have wide applicability often requires sacrificing some performance. In order to help address this issue during the conceptual design phase, it is useful to represent bearings as two lines with a line in between. Then after the rest of the concept is sketched, look and see where the point of intended action is in relation to where you placed the bearings, actuators, and sensing elements (if any). The more collocated the elements, the greater the degree of accuracy and control one is likely to have. Take the opportunity in the sketch phase of a concept to move the elements around in an attempt to minimize sine errors.

Consider the previous discussion of Saint-Venant's principle, where the product of the length of the bearing with the difference in the slope of the chassis and the axle potentially led to the loss of bearing clearance. This type of sine error is a frequent cause of machine element failure. Accordingly, one can often equate moments with angular deflections, and angular deflections with sine errors and the potential for misalignment and potential failure.

On the other hand, it is virtually impossible to design a structure without moments, so the goal should be to identify them and be aware of their consequences. An important strategy for dealing with them is to first reduce them (minimize the Abbe offset) whenever possible. Next make sure to assess their magnitude and effect. Finally, seek to use symmetry or one type of error motion to cancel another.

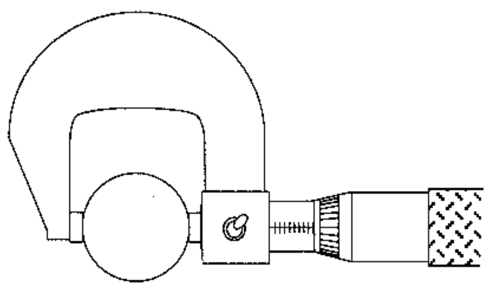
As an example, consider the common problem of aligning shafts that support a linear motion carriage. If the shafts diverge, they have horizontal parallelism error, along their length, the bearings will bind, unless a large clearance is provided; however, this causes a lot of lateral error motion in the system. On the other hand, if one of the bearings is mounted so it connects with a flexure to the carriage, as shown in the figure, then the horizontal parallelism is accommodated by the product of the distance of the carriage from the bearing and a small angular rotation of the bearing. Note that the amount of vertical error this introduces into the carriage's motion is only a cosine error. This principle was patented many years ago, and can now be used by anyone.

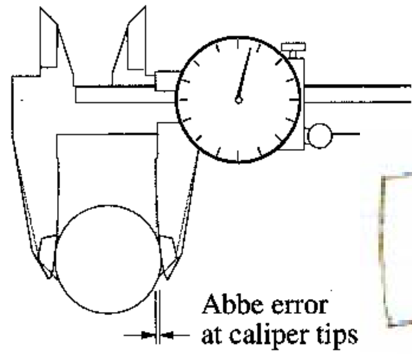
Sine errors can also be used to cancel other errors. An example is the alignment flexure that Dr. William Plummer of Polaroid Corp. created when he needed a fixture to adjust the pitch and roll of a lens without any accompanying translational motion. The vender proposed a complex set of nested angular motion cradles costing many thousands of dollars. Dr. Plummer created a system for a few hundred dollars by taking advantage of the fact that the slope at the end of a cantilever beam is proportional to $L^2/2$, whereas the deflection of the beam is proportional to $L^3/3$. If the lens were attached to an arm of length 2L/3 that was attached to the end of the beam, when the beam was deflected at its tip the lens would undergo pure tilt because the sine error would cancel the beam deflection error.

Look closely at your ***concepts*** and the requirements for precision. How might *sine errors* be used to your advantage to reduce system complexity or increase design robustness?

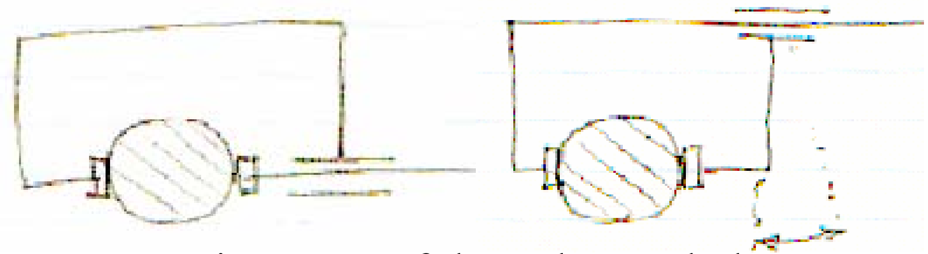
Abbe's Principle: *Locating Components*

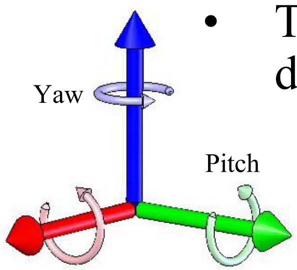
- Geometric: Angular errors are amplified by the distance from the source
  - Measure near the source, and move the bearings and actuator near the work!
- Thermal: Temperatures are harder to measure further from the source
  - Measure near the source!



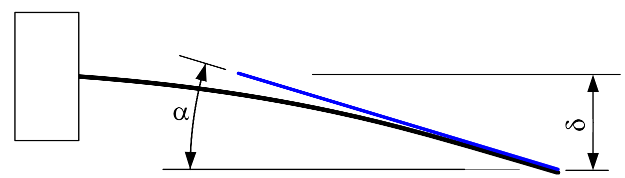


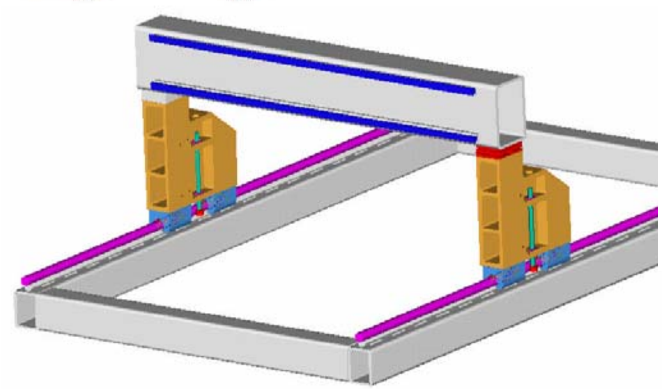
On Brown & Sharpe's vernier caliper: "It was the first practical tool for exact measurements which could be sold in any country at a price within the reach of the ordinary machinist, and its importance in the attainment of accuracy for fine work can hardly be overestimated"

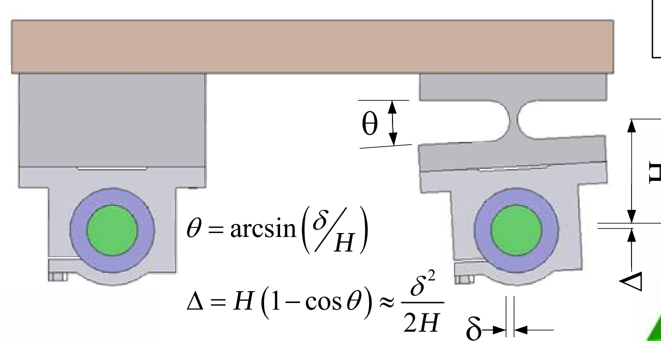


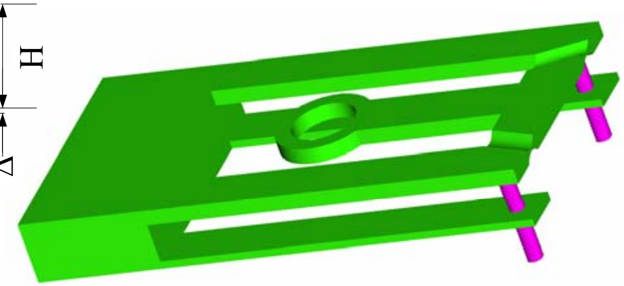


- Thinking of Abbe errors, and the system FRs is a powerful catalyst to help develop DPs, where location of motion axes is depicted schematically
  - Example: Stick figures with arrows indicating motions are a powerful simple means of depicting ***strategy*** or ***concepts***









Abbe's Principle: *Cascading Errors*

A single component and its angular deflections can lead to potentially significant sine or cosine errors. Many components together compound the problem. The expected performance can be conservatively estimated as the average of the sum of the errors and the root mean square of the errors. However, we must still ascertain the contribution of each of the errors of each of the ***components*** on the point of interest in the machine, given that the Abbe error or sine error is different for each component.

Consider the Lego™ elements shown which are made with amazing quality. Each brick must be a little bit smaller than the actual ideal size, because if each brick were a tiny bit larger, you could never assemble a wall of bricks. Thus each brick must be made smaller by an amount δ. Hence each brick in a wall can diverge from nominal linearity by δ/width-of-a-brick. The result of stacking together many bricks and then taking all the error out by pushing them all to one side is very dramatic as shown in the pictures!

The process of keeping track of all the errors in the system is called *error budgeting*. The goal of any machine is to position the "tool" on the "work" with an acceptable amount of error. Given that machines are complex 3D systems, the concept of sine and cosine errors needs to be considered in three dimensions. Thus to define the relative position of one rigid body with respect to another, six degrees of freedom must be specified. To further complicate matters, error in each of the six degrees of freedom can have numerous contributing components[1].

Considering all the interacting elements in a typical machine, the best way to keep track of and allocate allowable values for these errors is to use an error budget. An error budget, like any other budget, allocates resources (allowable amounts of error) among a machine's different components to help control and guide the design process and to help predict how the final design will behave. Like any budget, an error budget is a dynamic tool that must be continually updated during the design process. This is an advanced concept normally covered in graduate courses on precision machine design or kinematics; however, the basic concepts and can be practised by anyone.

1. A. Slocum, Precision Machine Design, 1994 SME Dearborn, MI USA.

To represent the relative position of a rigid body in three-dimensional space with respect to a given coordinate system, a 4x4 matrix is used which is called the *homogeneous transformation matrix* (HTM). This matrix represents the coordinate transformation to the coordinate system (X_{n-1} Y_{n-1} Z_{n-1}) from that of the rigid body frame (X_n Y_n Z_n). The first three columns of the HTM are direction cosines (unit vectors i, j, k) representing the orientation of the rigid body's X_n, Y_n, and Z_n axes with respect to the previous reference coordinate frame, and their scale factors are zero. The last column represents the position of the rigid body's coordinate system's origin with respect to the n-1 coordinate frame's origin. The pre-superscript represents the reference frame you want the result to be represented in, and the post-subscript represents the reference frame you are transferring from. The transformation matrices can be multiplied in series. The following transformation matrix implies that to move from reference frame *n-1* to coordinate frame *n*, one first moves along the X, Y, and Z axes of reference frame *n-1* by *a, b,* and *c* respectively, and then one rotates about the X axis by θ_x, and then in this rotated frame, about the Y axis by θ_y, and then in this rotated frame about the Z axis by θ_z and you are now in the reference frame *n* (where Sθ = sinθ and Cθ = cosθ):

$$
{}^{n-1}T_n = \begin{vmatrix} C\theta_y C\theta_z & -C\theta_y S\theta_z & S\theta_y & a \\ S\theta_x S\theta_y C\theta_z + C\theta_x S\theta_z & C\theta_x C\theta_z - S\theta_x S\theta_y S\theta_z & -S\theta_x C\theta_y & b \\ -C\theta_x S\theta_y C\theta_z + S\theta_x S\theta_z & S\theta_x C\theta_z + C\theta_x S\theta_y S\theta_z & C\theta_x C\theta_y & c \\ 0 & 0 & 0 & 1 \end{vmatrix}
$$

Given a position vector $[d\ e\ f\ 1]^T$ representing the x, y, z coordinates of a point in coordinate system n, then the coordinates of this point in coordinate system n-1 will be: ${}^{n-1}T_n[d\ e\ f\ 1]^T$. In this manner, the position of the "tool" in a reference frame can be compared to the position of the "work" in the same reference frame, where all the errors in all the coordinate frames have been taken into account using all the HTMs between all the individual elements in the machine. The spreadsheet *ErrorGainSpreadsheet.xls* helps to make the use of HTMs much easier (see pages 10-35-10-38).

How would you account and allow for all the errors between all of the ***components*** in your ***concepts*** to ensure the system will work as desired?

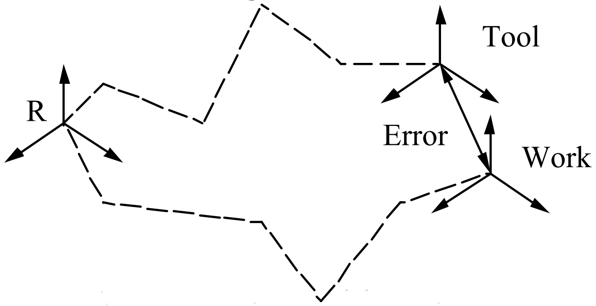
Abbe's Principle: *Cascading Errors*

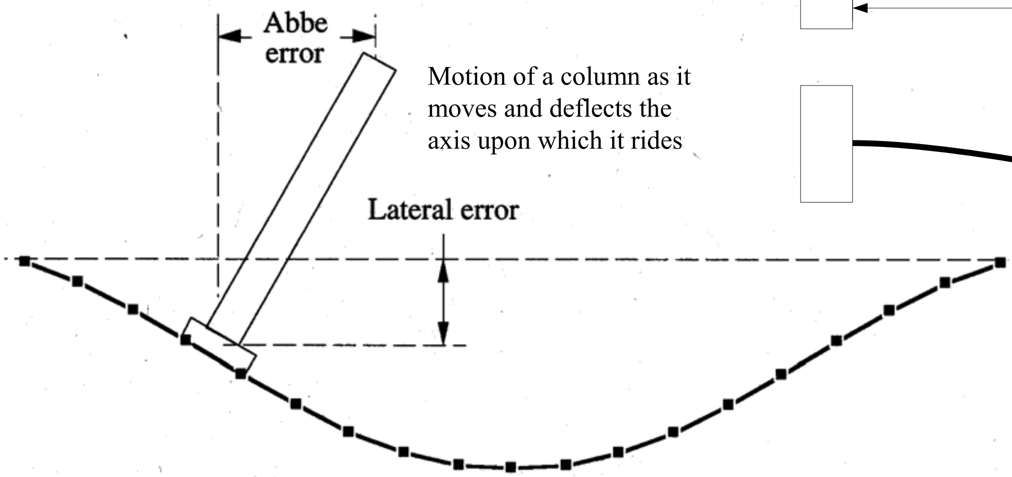
- A small angular deflection in one part of a machine quickly grows as subsequent layers of machine are stacked upon it…
  - A component that tips on top of a component that tips…
  - If *You Give a Mouse a Cookie*… (great kid's book for adults!)
- Error budgeting keeps tracks of errors in cascaded components
  - Designs must consider not only linear deflections, but angular deflections and their resulting *sine errors*…

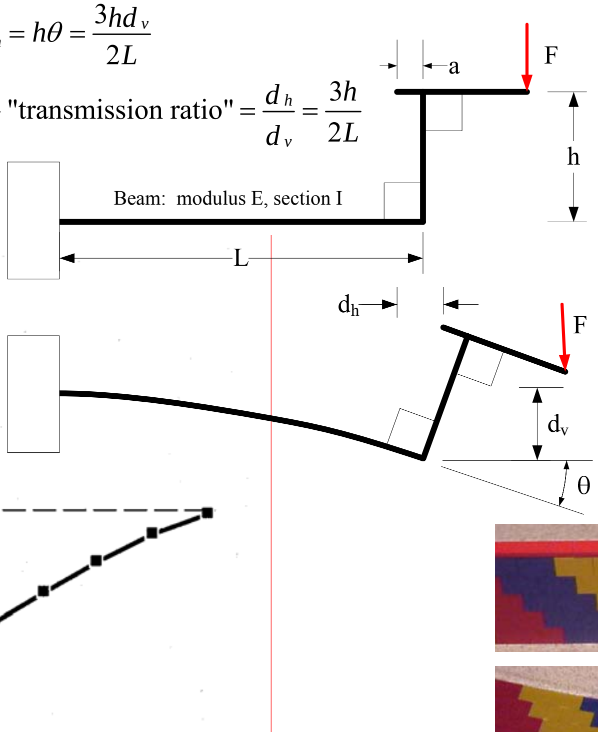
$$d_v = \frac{FL^3}{3EI} \qquad \theta = \frac{FL^2}{2EI} = \frac{3d_v}{2L}$$

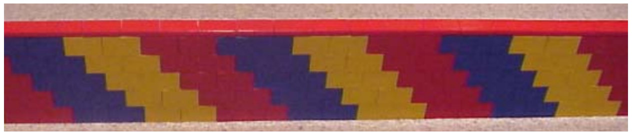
$$d_h = h\theta = \frac{3hd_v}{2L}$$

$$\Rightarrow \text{"transmission ratio"} = \frac{d_h}{d_v} = \frac{3h}{2L}$$









Maxwell & Reciprocity

Yin and Yang, and *do unto others as you would have others do unto you*... have implications for how many people live their lives, and also interestingly enough have a strong parallel in the engineering world in the form of Maxwell's principle of *Reciprocity*. Like many geniuses of the 19th century, James Clerk Maxwell was an applied mathematician, and mathematics knows no bounds; thus he was as comfortable in the world of mechanics as he was in the world of electromagnetics. As defined, his theory of reciprocity applies to mechanics and forms the basis of modal analysis which is critical to the dynamic evaluation of high performance systems. In addition, the principle of reciprocity applies to precision measurement systems and philosophically to how difficult problems can be creatively solved.

Reciprocity is commonly applied in mechanics to understand the dynamic behavior of structures and machines, and is thus of critical importance to machine designers. *Experimental Modal Analysis*[1] allows the study of vibration modes in a machine or structure. An understanding of data acquisition, signal processing, and vibration theory is necessary to obtain meaningful results, but it is easy to understand what the process yields and philosophically how it is done. The results of a modal analysis yield modal natural frequencies, modal damping factors, and vibration mode shapes This information may be used to: Locate sources of compliance in a structure, characterize machine performance, optimize design parameters, identify the weak links in a structure for design optimization, identify modes which are being excited by the process (e.g., an end mill) so the structure can be modified accordingly, and identify modes (parts of the structure) which limit the speed of operation.

Experimental modal analysis requires measurement of input and output signals to and from a system using appropriate transducers and analog to digital converters. The input is usually a force excitation. The output may be measured with an interferometer, a capacitance probe, an accelerometer, or another response transducer. Many machines may be conveniently analyzed with inexpensive piezoelectric force and acceleration sensors. A multi channel dynamic signal analyzer is used to obtain good quality time histories from the signals.

1. This section is distilled from lecture notes created by Prof. Eric Marsh (ERM7@PSU.EDU) of Penn State when he was one of Prof. Slocum's graduate students and they used to help companies find and fix vibration problems.

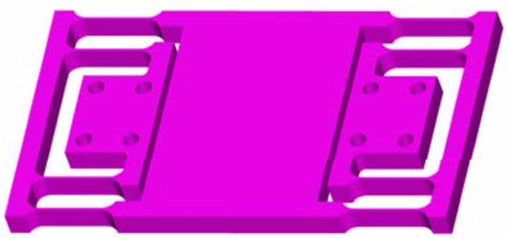
A *Fast Fourier Transform* is used on the discrete time data to obtain the Frequency Response Function (FRF) between input and output. Calculation of input and output FFT's allows the computation of the transfer function, which is evaluated along the jw-axis and therefore called the frequency response function (frf). The coherence, which gives an indication of the quality of the data, may also be calculated. A 0 indicates poor quality, 1 indicates high quality. The frf for that point on the machine is then stored on disk.

The process of collecting and storing the FRFs is repeated over many points on the structure. Either the location of the input or the output measurement point is changed, reciprocity assures us it does not matter, as long as one is consistent: Either channel, but not both, may be moved as a result of reciprocity in linear systems. An entire data set is thus collected by repeating the measurement process over many locations on the test article.

This collection of FRFs is used to determine a machine's natural frequencies and modal damping factors. The collected FRFs will show the same modes of vibration; each FRF will have peaks at the same frequencies with the same amount of damping. The difference will be in the magnitude of each peak. The drive point, which is typically made at the point of most interest on the structure, such as a spindle of a machine tool FRF is typically a good FRF to use for locating the modal frequencies and damping factors.

For each mode, measure fluctuation of response amplitude over all the collected FRFs. Each FRF is now used to estimate the mode shapes of vibration. The magnitude of each vibration mode is recorded for each of the collected FRFs. A large magnitude for a given mode in a given FRF indicates that the structure has a large amplitude at that location and frequency (antinode). Small magnitudes indicate that the structure is barely moving at the indicated location and frequency (node). The software then uses magnitudes of the mode shapes to animate a wireframe mesh on a computer. This helps visualize each vibration mode and identify sources of compliance in the machine, and regions lacking damping.

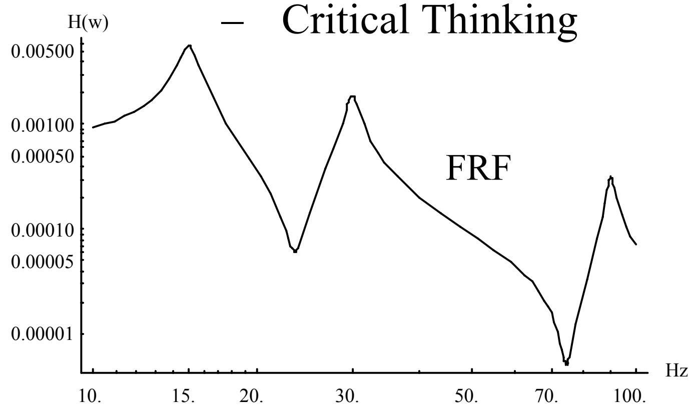
Look at your ***concepts*** and envision what happens as they speed up. How will the motors and actuators respond as they speed up? If you were to scale your machine up into a machine that you could drive in, would it ride as smooth as a good car? How might you increase your machine's dynamic performance?



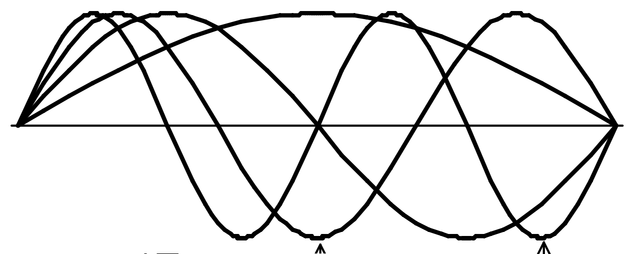
Maxwell & Reciprocity

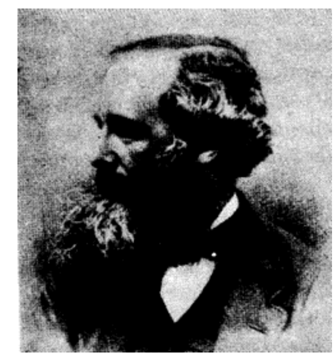
$$\frac{1}{problem} = opportunity! \quad \frac{1}{Ow!} = Ahhhhh!$$

- Maxwell's theory of *Reciprocity*
  - Let *A* and *B* be any two points of an elastic system. Let the displacement of *B* in any direction *U* due to a force *P* acting in any direction *V* at *A* be *u*; and the displacement of *A* in the direction *V* due to a force *Q* acting in the direction *U* at *B* be *v*. Then $Pv = Qu$ (from Roark and Young Formulas for Stress and Strain)
- The principle of *reciprocity* can be extended in philosophical terms to have a profound effect on measurement and development of concepts
  - Reversal
  - Critical Thinking

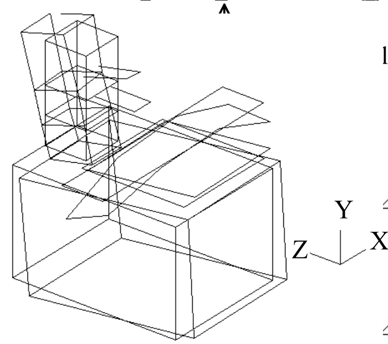


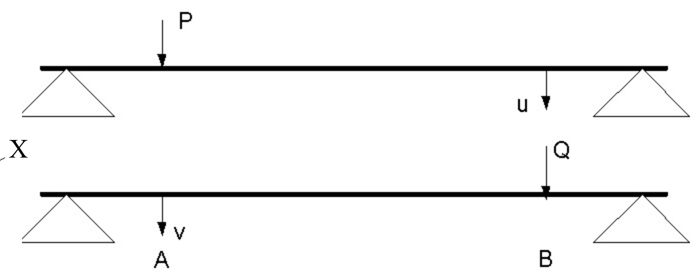
Drive point measurement

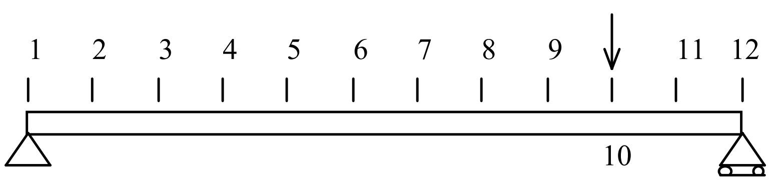




James Clerk Maxwell 1831-1879







Maxwell & Reciprocity: *Reversal*

Have you ever wondered how machine tools were made in the first place? Who made the machine that made the machine in the shop have such great accuracy? It all started with the principle of reciprocity long before Maxwell described the formal mathematics. Starting with three square plates, A, B, C, the surface of each plate is rubbed with a colored compound (e.g., rouge) and then rubbed on one of the other plates. High points rub and are thus exposed, and these high points can then be scraped off by hand. The process repeats in a round-robin fashion until the desired flatness is achieved. Note that in order to reduce twist of the surface, square plates must be used, so they can also be rotated 90 degrees as the round-robin process repeats. From a flat surface, a square cross section straightedge can be made. One square straightedge can be placed next to another on a surface plate, and then right angles achieved by rubbing and scraping[1]. The best surfaces are called *masters*, and from these, other references surfaces can be made.

The next step is to divide a circle. It is well known how to bisect an angle with a compass, and starting with this process, a circle can be divided. Similarly, a length standard can also be divided. From these basic methods, standards evolved which in turn gave birth to the modern world. However, it was not until the 1900s that reciprocity was really put to the test, with the creation of ruling engines to make diffraction gratings[2]. Ruling engines used leadscrews to move an aluminum coated blank under an oscillating diamond tip which furrowed straight paths into the surface. Accuracy and line density depended on one turn of the screw creating the same distance travelled anywhere along the length of the screw. This was near impossible to achieve, so the errors in screw pitch were compensated for by a cam which caused the nut to rotate slightly as the screw was turned to compensate for screw pitch errors. These corrector cams allowed gratings to be ruled with 25,000 lines per inch (one micron pitch)! It was not until the 1930s that Prof. George Harrison at MIT used a laser interferometer and electronic feedback to rule accurate gratings with over 100,000 lines per inch! So you see, from 3 simple plates and reciprocity, we get the machines that made the instruments that launched our modern world.

1. An excellent description of the process is given by Ted Busch Fundamentals of Dimensional metrology, Delmar Publishers, Inc., 1966, Albany, NY. Also see Topic
2. Recall that diffraction gratings can be used to diffract an incoming beam and separate it into its component wavelengths; thus they are useful from studying everything from atoms to the cosmos.

Reversal was thus in use for hundreds of years, but in the 1970s Dr. John Simpson and Dr. Robert Hocken of the National Bureau of Standards[3] used the reversal principle to form the basis of electronic error correction of Coordinate Measuring Machines (CMMs) and machine tools. They realized that machines can more easily be made mechanically repeatable than they can be made accurate; thus they hypothesized that the errors could be measured by having the machines measure known accurate artifacts, and the errors in measurement as a function of the machine's axes' positions could be recorded and compensated for in the machine's controller.

The most accurate artifacts are needed to act as standards to which all other manufactured parts can be compared. Artifacts themselves were made using the reversal process. As illustrated, an artifact of unknown straightness is measured using a CMM. The artifact is then flipped over (reversed), measured again, and the two records of error as a function of distance along the artifact are subtracted from each other. The reversal causes the error in the part, as long as axial position has not changed, to change sign, while the contribution to the error of the CMM's axes remains the same; hence repeatable errors in the CMM's motion are subtracted out and we are left with twice the part error. Repeatability of the CMM itself can be ascertained by using a single probe held at a single position, and the machine axes can be moved back and forth, so how well the machine repeats its position can be measured.[4]

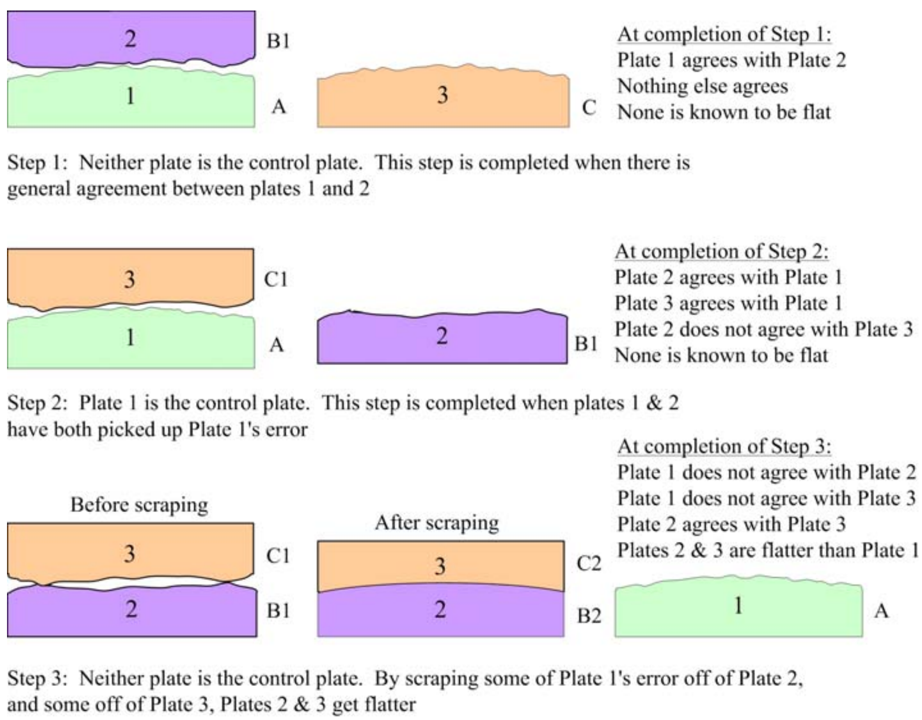
Reversal is also useful in manufacturing. From flipping the grain on boards to be glued in wood shop, to using the bow in one surface to cancel the bow in another, reciprocity is one of the most powerful fundamental principles in the design engineer's toolkit. If it is used in the design of machines themselves, more robust machines can be created in the first place. This is the essence of *self-help discussed on page 3-14*.

Review your ***concepts*** and ***strategies*** and identify what might be considered weak points or potentials for performance reducing inaccuracies. How might reversal be used to compensate?

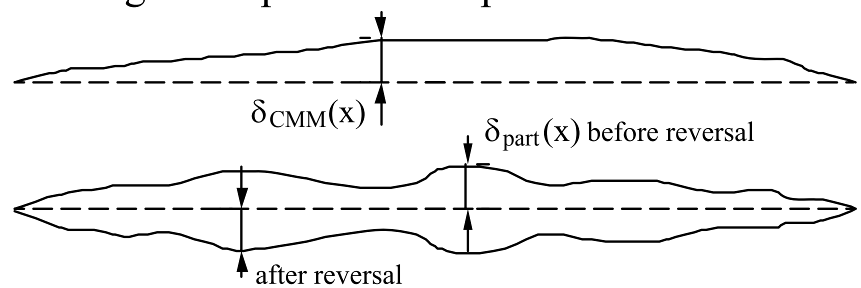
3. Now the National Institute of Standards and Technology (NIST).
4. Dr. Robert Donaldson of Lawrence Livermore national laboratory was one of the pioneers of reversal methods. These are discussed in great detail in the ANSI *Axis of Rotation* Standard B89.3.4

Maxwell & Reciprocity: *Reversal*

- *Reversal* is a method used to remove repeatable measuring instrument errors
  - A principal method for continual advances in the accuracy of mechanical components
- There are many applications for measurement and manufacturing
  - Two bearings rails ground side-by-side can be installed end-to-end
    - A carriage whose bearings are spaced one rail segment apart will not pitch or roll
  - Scraping three plates flat



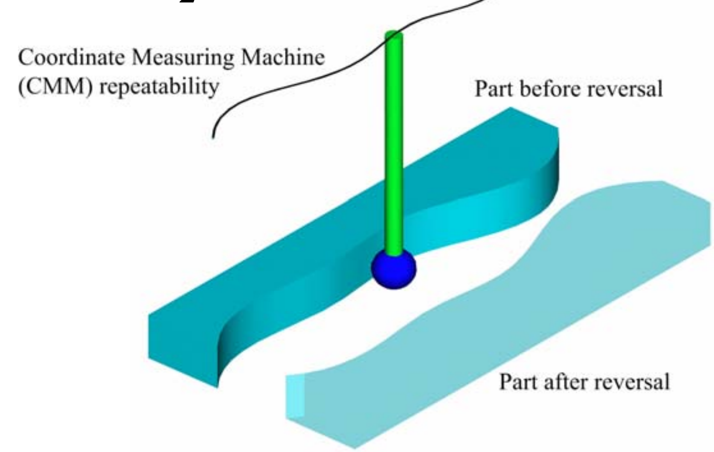
After T. Busch, Fundamentals of Dimensional Metrology, Delmar Publishers, Albany, NY, 1964



$$Z_{\text{probe before reversal}}(x) = \delta_{CMM}(x) - \delta_{part}(x)$$

$$Z_{\text{probe after reversal}}(x) = \delta_{CMM}(x) + \delta_{part}(x)$$

$$\delta_{part}(x) = \frac{-Z_{\text{probe before reversal}}(x) + Z_{\text{probe after reversal}}(x)}{2}$$



Maxwell & Reciprocity: *Critical Thinking*

Can reciprocity be applied to the creative process? ***There should be no problems, only opportunities.*** A simple way to learn to use reciprocity in critical thinking is to play word association games with words that describe problems you encounter in everyday life, as well as with problems encountered during the design process[1]. Consider your own personal comfort: If you are uncomfortable on your back, *roll over on your stomach.* If too many clothes make you too hot, *take some off.* If you are board skiing, *try snowboarding!*

An effective way to apply reciprocity to critical thinking is to have a clear understanding of the physics of the problem, both in philosophical and analytical terms. What are the *dominant parameters*, or the most *sensitive directions.* Controlling or inverting sensitive directions by adding or changing the design can have the greatest benefit. Consider being too hot or too cold. Yes, it is obvious to try and use clothing to regulate your temperature, but which articles of clothing? One has to look at the body and see what are the most heat sensitive zones. The head and neck are regions with the most surface vascularity, and blood flow is not constricted to the head when you are cold, so regulating head coverings is the most effective means of controlling body temperature. Your hands and feet are also very vascular and also have a high concentration of sweat glands. Now consider the design of hats and gloves and socks and shoes, and the general binary state of their design...

One day while driving, I was caught in a huge traffic jam. A truck carrying a roll-off container had an accident and the container flipped-off the truck and crushed a car. Despite regulations requiring containers to be chained down, drivers never seem to do it because it takes time, they forget... What is needed is a passive system. However, when I looked under a roll-off container on a truck, I realized there was no space for a strong-enough attachment mechanism: If only the forces that caused the moments could be used to engage and hold elements together, rather than pry them apart. My friend Bill Miskoe made the parts and another friend allowed them to be welded to one of his trucks. Bill then used his crane to pick up and shake a truck by the container. It worked great!. Bill and I then went to various trucking and container manufacturing companies. No one was interested because a) the handful of people killed every year was not worth bothering with (That is what they had insurance for), or b) they had their own system that just required the driver to attach a simple chain.... maybe lawyers are not so bad after all.

Consider designing injection molded plastic snap-fits. By reversing the position of the male and female snap-fit elements, the need for a side pull can sometimes be replaced with a simple protrusion on each of the simple mold elements.

If clearance between elements allows one to move relative to another and cause noise or unwanted motion, *use an element in the opposite direction to cancel the unwanted motion*. *Backlash* is clearance between the threads of a nut and a threaded shaft's threads which must exist or the threads will jam, but this makes the position of the nut uncertain. Making the threads with little or no clearance is very expensive and the effect is lost as the threads wear. Instead, consider a second set of threads which pushes against the first nut by means of a spring. The threads are always in contact regardless of the direction of travel. The second set of threads can be another nut and the spring can be a coil spring, spring washer, or a flexure. A single nut can also be split longitudinally and then a circumferential spring applied to cause radial preload of the threads...

Like being too hot or too cold, a system can be too stiff or too compliant. The answer may not be to use bigger or smaller elements, it may be to use many instead of few, or few instead of many. For example, if a part will not rest flat on a surface so it wobbles, can it be instead supported by three points? Or can a compliant layer be placed under the part, so in effect the part is supported by millions of tiny supports, each able to easily deflect to accommodate local imperfections, yet together they form a very rigid support? An example is bonding metal parts together, where a thin adhesive layer does not result in significant loss of stiffness because even though the modulus of the glue is far less than the metal, because the compliance contribution of the adhesive layer is proportional to its low modulus divided by its even lower thickness.

Carefully review the risks associated with your ***strategies*** and ***concepts***, and see if risks can be mitigated by inverting them either literally, or using analysis to identify the most sensitive parameters which then lead to new or modified mechanism or ***strategies***.

1. An interesting reciprocity website is www.whynot.net

Maxwell & Reciprocity: *Critical Thinking*

Rock & Roll Over & Under

- If you are:
  - Happy, turn it around!
  - Unhappy, turn it around!
  - Comfortable on your back, turn over and try lying down on your front……
- You can make a system *insensitive* to its surroundings, or you can *isolate* a system …
- If you cannot solve a problem by starting at the beginning, work backwards!
- Example: Roll-off container passive restraint mechanism
  - In the event of an accident, it keeps an otherwise gravity-held container from flipping off the truck

Reciprocity
It's like velocity
Once up to speed
You have no other need
Late at night
On goes the light
Driven by curiosity
You create with ferocity

Bill Miskoe, welder and co-inventor

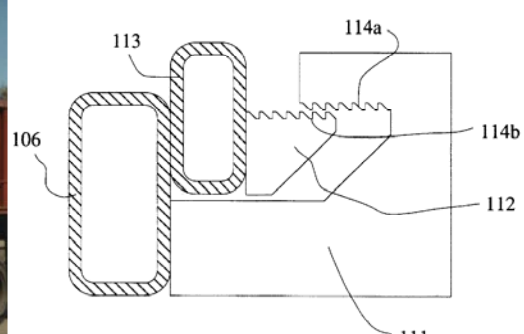


Fig. 7

United States Patent [19]
Slocum et al.

[11] **Patent Number:** 5,848,869
[45] **Date of Patent:** Dec. 15, 1998

[54] **CONTAINER RESTRAINING MECHANISM AND METHOD**

[75] Inventors: **Alexander H. Slocum**, Bow; **John William Meskoe**, Concord, both of N.H.

[73] Assignee: **AESOP, Inc.**, Concord, N.H.

[21] Appl. No.: **759,870**

[22] Filed: **Dec. 3, 1996**

[51] **Int. Cl.**[6] **B60P 1/65**
[52] **U.S. Cl.** **414/500**; 414/480; 414/494
[58] **Field of Search** 220/1.5; 414/480, 414/491, 492, 493, 494, 499, 498, 500

[56] **References Cited**

U.S. PATENT DOCUMENTS

3,897,882	8/1975	Budoff	414/500
5,085,448	2/1992	Shubin	414/500 X
5,284,266	2/1994	Januel et al.	414/498 X

FOREIGN PATENT DOCUMENTS

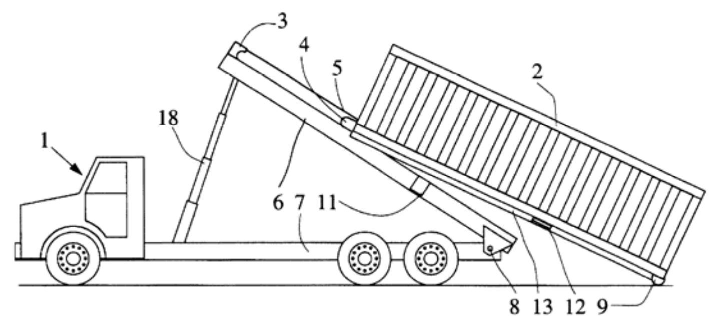
71143	1/1976	Australia	414/500
481367	4/1992	European Pat. Off.	414/480
1080210	12/1954	France	414/500
2686843	8/1993	France	414/500
1430217	3/1969	Germany	414/494
8607019	12/1986	WIPO	414/494

Primary Examiner—Frank E. Werner
Attorney, Agent, or Firm—Rines & Rines

[57] **ABSTRACT**

A system and technique for holding down and restraining a roll-off container in place as it is winched into its final position on the bed of a truck or other transport device, wherein, as the roll-off container is winched into position on the truck, a protuberance on the side of the bottom rail of the roll-off container slides into a mating interlocking cradle attached to the truck. The sliding mate can be of a cantilevered spear into a socket, or more generically, a sliding open-ended mate such as a dovetail. In the event of a crash or sudden stop, the roll-off container will thus be retained by the interlocking connection.

5 Claims, 4 Drawing Sheets



Self-Principles

The principle of reciprocity leads naturally into the principle of using the problem or structure itself as the means of its solution. The technique is simple, just say "I have a problem with" and replace this phrase with "Self-". For example, can you envision systems which have need of the capability for *Self-Reinforcing, Self Balancing, Self-Limiting, Self-Protecting, Self-Damaging, Self-Braking, Self-Starting, Self-Releasing, Self Serving...?* To help achieve each of these effects, one of the strongest starting points in design is to use the principle of *Self-Help*, which is essentially the *principle of reciprocity.*

As an example of *Self-Reinforcing* or *Self-Help*, consider early in the industrial revolution when boilers started to proliferate to generate steam for power. It became apparent that the boilers needed to be cleaned to prevent buildup of crud which reduced efficiency and also led to corrosion with disastrous side effects; however, to make a boiler, it had to be riveted tightly together so there would be no leaks. To put an access door in the structure would be to invite leaks from the internal pressure which was sure to deform any door so the steam could escape. Somewhere, someone thought to make the door like a tapered plug, so the higher the pressure, the more the door will deform to conform to the surface against which it must seal. Airplane doors work under the same principle, so when the plane is flying high, and the differential pressure between the inside and outside of the plane my be 1/4 atm., the force holding the door shut may be on the order of 50,000 N!

As an example of *Self Balancing*, consider your washing machine, where it is near impossible to ensure that the clothes will be properly distributed to obtain dynamic balance during the spin cycle. Somewhere, someone figured that if elements where placed in the drum that were free to move, they could be designed to move to the proper place in the presence of dynamic imbalance; hence was born the three-ball balancer which uses three balls free to move in a groove along the outside of the drum.

An example of *Self-Limiting* is the flyball governor, which is what James Watt really invented in regards to making steam engines. As the speed of the engine increases, which would cause centrifugal stresses to become too great in many parts of the engine, centrifugal forces cause balls attached to levers to move outward, and the levers gradually close a valve controlling steam flow; hence velocity is controlled.

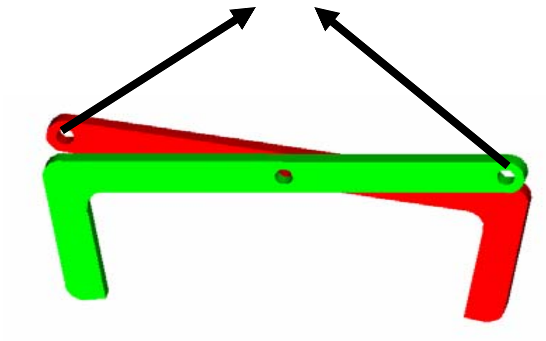
As another example of *Self-Limiting*, consider turbine blades where tangential forces from the work they do on the fluid cause bending stresses at the root of the blades. Also consider the stresses caused by centrifugal forces. Tensile stresses are the real problem, because they cause cracks to grow. Bending stresses cause tensile and compressive stresses. Can the blades' center of mass be located such that as the blade spins faster the bending moments are opposed to each other, so tensile stresses are reduced?

An example of *Self-Protecting* is to counter what you are afraid of by learning to do it. If you are afraid of falling, learn how to fall (take judo!). If you are afraid of being attacked, learn how to attack If you are afraid of love, then learn to love. If you are afraid of a current or voltage surge burning out your system, install a circuit breaker or fuse where the surge trips the breaker and shuts itself off. If you are afraid of lightening striking your house, encourage it to strike where it will do no damage by installing lightening rods.

An example of *Self-Damaging* are price labels that are applied with microcuts in them, so they look fine, but when they are removed, they come apart and thus are very hard to pull off. People who would change the price tags in stores before check out hate them! What about stickers that cannot be removed without coming apart, thereby clearly showing if a customer tried to service non-customer serviceable components? What about some things people do to themselves in the name of fun or productivity?

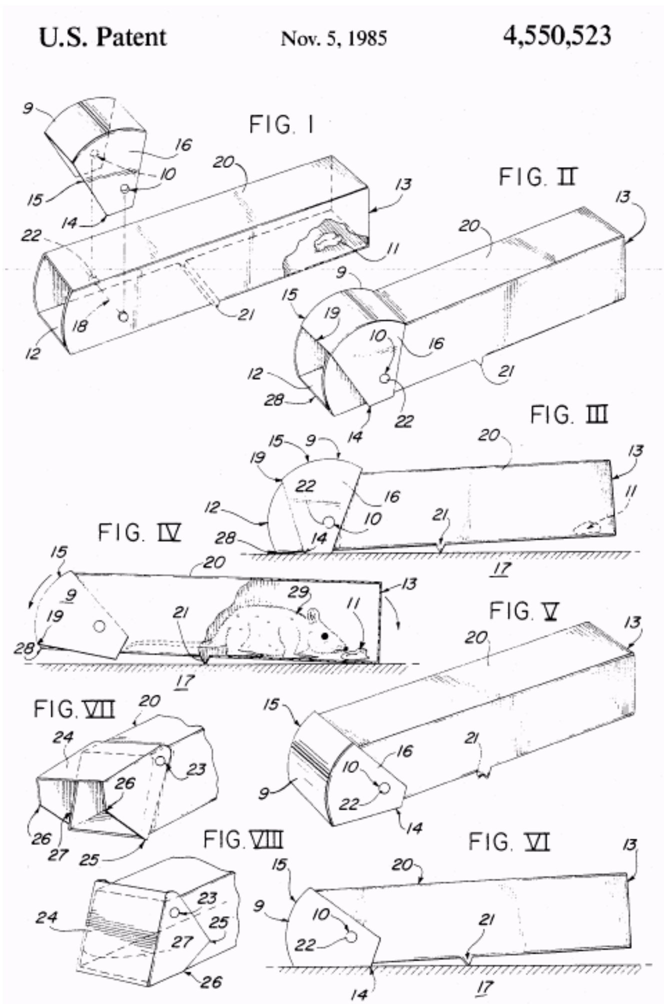
As an example of *Self-Braking*, consider the need to rapidly stop an object. If a tapered brake element is gently pushed into the system such that frictional forces pull the tapered element in even harder, then the system will stop itself very quickly. An example of *Self-Starting* is an avalanche, or more constructively doing your homework before your dad has to remind you, or using a hill to get your old car with a dead battery up to speed before popping the clutch...A system that is *Self-Releasing* can use the rate of a process to cause a latch to release, such as the burning of fuel in a rocket triggering explosive bolts to release a load.

What contest processes must be initiated, or controlled in your ***strategy***, and how can your ***concept*** handle them without you having to control them? Can ***strategies*** you discard be countered if others decide to use them?



Self-Principles

- The manner in which a design reacts to inputs determines its output
  - Reciprocity would philosophically tell us to look for a solution where a potentially detrimental result can be used to cancel the effect
  - Martial artists practice this principle all the time!
- *Self-Help:* A design that uses the inputs to assist in achieving the desired output
  - An initial effect is used to make the device ready for inputs
    - The supplementary effect is that which is induced by the inputs, and it enhances the output
  - *Example:* Airplane doors act like tapered plugs
    - When the door is shut, latches squeeze the seal, making the cabin airtight
    - As the plane ascends and outside air pressure decreases, the higher inner air pressure causes the door to seal even tighter
  - Example: *Back-to-back* angular contact bearings are thermally stable
  - Example: Ice tongs
  - Example: A better mousetrap!
  - Example: Balanced forces on hydrostatic bearings: A.M. van der Wielen, P.H.J. Schellekens, F.T.M. Jaartsveld, Accurate Tool Height Control by Bearing gap Adjustment, Annals of the CIRP, 51(1/200), 351-354, (2002)
- Other *self-principles* similarly exist:
  - *Self Balancing, Self-Reinforcing, Self-Protecting, Self-Limiting, Self-Damaging, Self-Braking, Self-Starting….*



Stability

The principles of *reciprocity* and *self-help* are naturally applicable to, or perhaps catalyzed by, the consideration of the state of *stability* in a system. However, in order to design a system to be stable, one must understand what makes it unstable! For some systems, such as bulk material feeders or compactors, induced vibration, caused for example by a rotating eccentric mass, can be essential. Stability transitions can also be exploited. A common example is a snap-fit that uses an applied force to move a part from a *stable*, to an *unstable*, and finally to a new stable position. Stability is of utmost importance in the design of structures, actuators, bearings, and control systems.

In the case of structures, stability is defined in terms of not breaking, but there are specific types of stability issues associated with members in compression, such as truss elements. As the compressive load on a long slender member increases, infinitesimal lateral deflections are acted on by the axial force to become bending moments, which increase the deflections and the moments so the member eventually buckles[1]. The buckling force is a function of the material modulus, moment of inertia, length, and a constant C that depends on the mounting:

$$F_{Buckle} = C\frac{EI}{L^2} \qquad \omega_{rpm} = \frac{60k^2}{2\pi}\sqrt{\frac{EI}{A\rho L^4}}$$

Actuators, which often have rotating components, such as spinning shafts, experience a similar effect. If the shaft is not exactly 100% perfect, as it spins, the imperfection creates a centrifugal force which causes the shaft to become unstable when the speed reaches the first natural frequency of the shaft. This is called *shaft whip*. In some systems, such as some high performance turbine systems, the turbine can run at super critical speeds by accelerating through the resonance point. The system just must not ever operate at or near resonance.

As an example of using the principle of *stability* to make a system *self-limiting*, consider a machine whose axes are actuated with a leadscrew. As the screw turns, the nut moves, and so the shaft is constantly seeing changing boundary conditions. Conventional conservative designers would limit the maximum screw shaft speed to the speed that causes shaft whip when the nut is all the way at one end of travel. On the other hand, a wise design engineer will program the control system to keep the shaft speed below the critical speed which is actually a function of the nut position along the shaft.

1. This potential bad thing can be made a good thing, as Minus K Technology has found, because it can be used to decrease the stiffness of a flexure thereby making a better flexural bearing. See www.minusk.com

Recall the discussion of Saint-Venant’s principle with respect to bearing design to reduce the chance of sliding instability (e.g., a drawer jamming). Instability occurs as the width/length spacing ratio of the bearings decrease in the presence of finite friction, which causes the bearing reaction forces to rise, which causes greater deflections, which can cause the drawer to yaw (twist) and suddenly jam. The lesson is that stability has different degrees.

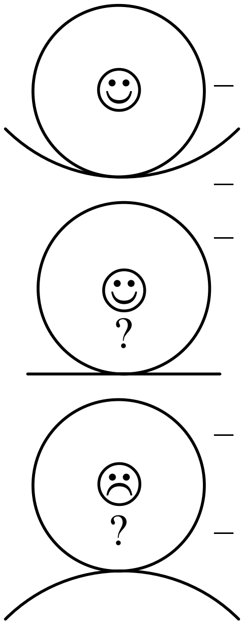
One of the most profound affects of stability on bearings is that how the bearing is mounted affects how the loads on the bearing change as the bearing heats up due to viscous shear in the bearing lubricant at high speed. Bearings mounted in the *back-to-back* mode use axial thermal growth of the inner ring to cancel radial thermal growth and thus remain thermally stable at high speeds. This mounting is also very resistant to moment loads, but this also makes it intolerant of misalignment. Bearings mounted in the face-to-face mode have completely opposite properties. They are tolerant of misalignment, but not thermally stable. Bearings and are discussed in great detail in Chapter 10.

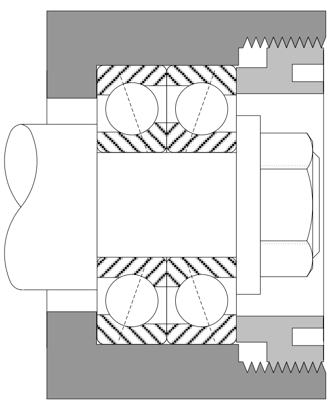
With regard to control systems, some systems are stable under an extremely wide range of conditions, and thus they are very robust. Some systems are at best stable only with very careful control. However, many such systems are very high performance, and as one learns in the study of advanced control systems, marginal stability can be exploited to achieve high speed performance. An example is the forward-swept wing fighter plane, which is only made stable by computers constantly controlling the flight surfaces, yet it can turn far faster than any other type of aircraft. Another example is walking or running, which are really controlled falling!

What sort of contest table *stability* issues can be exploited to minimize the effort of scoring? What are the *stability* trade-offs in your machine? How can a *stability* transition be used to your advantage?

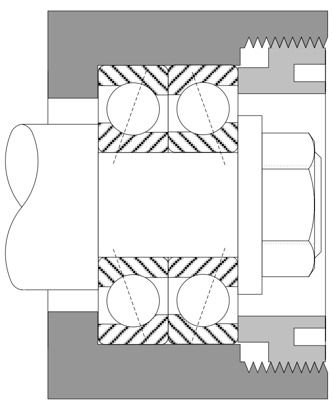
Stability

- All systems are either *stable, neutral, or unstable*
  - Saint-Venant's principle was applied to bearing design to reduce the chance of sliding instability (e.g., a drawer jamming)
  - A snap-fit uses an applied force to move from a stable, to a neutrally stable, to an unstable to a final new stable position
  - Wheels allow a system to roll along a flat surface
  - As the load on a tall column increases, infinitesimal lateral deflections are acted on by the axial force to become bending moments, which increase the deflections….
    - Reciprocity says this detrimental effect can be useful: fire sprinklers are activated by a column that buckles when it becomes soft…
  - *Back-to-back* mounted bearings are intolerant of misalignment, but use axial thermal growth to cancel radial thermal growth for constant preload and thermal stability at high speeds
  - *Face-to-face* mounted bearings are tolerant of misalignment, but axial thermal growth adds to radial thermal growth and causes the bearings to become overloaded and seize at high speeds



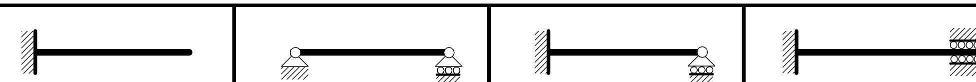


face-to-face mounting *can* accommodate shaft misalignment but *cannot* tolerate thermal expansion at high speeds



Back-to-back mounting *cannot* accommodate shaft misalignment but *can* tolerate thermal expansion at high speeds

$$\omega_n = k^2\sqrt{\frac{EI}{\rho L^4}} \qquad F_{buckle} = \frac{cEI}{L^2}$$



	Cantilevered		Simply Supported		Fixed-Simple		Fixed-Fixed	
mode n	k	c	k	c	k	c	k	c
1	1.875	2.47	3.142	9.87	3.927	20.2	4.730	39.5
2	4.694		6.283		7.069		7.853	
3	7.855		9.425		10.210		10.996	
4	10.996		12.566		13.352		14.137	
n	(2n-1)π/2		nπ		(4n+1)π/4		(2n+1)π/2	

Symmetry

Why do we have two arms, two legs, two eyes, two lungs...and only one nose, one liver...? We may seem symmetric, yet their are exceptions which are ultimately related to system cost issues. In general, symmetry is a good thing in a system, and thus is a good starting point for a design. Often a principle advantage of a symmetric system is that the part count is reduced, and the system becomes easier to manufacture, but not always! (see page 1-15).

Symmetry is a principle that is closely related to stability. Symmetry can sometimes provide stability, such as when a composite structure is made symmetric so differential thermal expansions balance each other. Moments are not created and the system only expands, but does not bend. Recall in the discussion of Abbe's principle that bending creates angular deformations which are amplified with distance, and hence are generally to be avoided. An example that might be encountered by making high precision machines from modular components would be the attachment of steel bearings to an aluminum structure. The bearings can be attached at the system neutral axis, or plain steel elements of the same cross section as the bearing rails can be attached to the opposite side of the structure to make it symmetric and thermally stable.

Symmetry can also sometimes lead to instability in highly dynamic systems. There is the classic example of the Tacoma Narrows bridge which did not owe its instability specifically to a symmetric design, but it was a contributing factor. Consider the marching of troops across bridges in WW II, where maintaining in-step formation could cause bridges to fail. What about automobile tires? If the tread pattern was circumferentially symmetric, very undesirable excitations could be created.

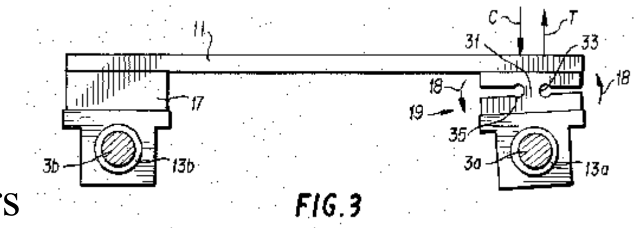
In structures, symmetry may exist on an overall scale, such as between two sides of a machine, but within each side of the machine, symmetry takes second place to the desired formation of triangles, real or virtual, so as to make a system well-braced and resistant to shearing. One of the main reasons a structure is often symmetric, is due to the fact that it must take loads from either side.

In bearings, it is often the opposite, where on an overall scale one side of the system has bearings configured and mounted to restrain five degrees of freedom of a shaft, and at the other end of the shaft, the bearings only prevent radial motion of the shaft. If the design where symmetric, with bearings on both ends of the shaft preventing radial and axial motion, then as the system temperature changed, the axial expansion of the shaft could create huge forces that would either cause the bearings to fail, or the shaft to buckle. The exception to this, is in ballscrews where the bearings on both ends are sometimes used to pre-tension the shaft, so thermal strains replace mechanical strains and the system not only remains stable, it can achieve a fixed-fixed mounting configuration, and hence achieve higher speeds. This is an advanced concept discussed in greater detail in Topic 6.

Symmetry can also be a two-edged sword with respect to the mounting of linear motion bearings. This is due to the fact that two bearing rails can never be made to be exactly parallel, and thus one bearing must be the master, and the other the follower. The exception is if compliance or clearance is provided such that the amount of rail non-parallelism can be accommodated in the bearings without overloading them. In general, the two bearing rails are bolted in place with as high parallelism as possible, and then the bearing blocks on one side are positioned against a reference edge on one side of the carriage, thus establishing the carriage parallel to the bearing rails. The bearing blocks on the other rail are then merely bolted to the carriage. If they were forced to a reference edge on the carriage, the system would be over constrained and the bearings would likely be overloaded.

In the history of machine tools, the first machine tools used a vee-shaped bearing rail to guide the motion of a carriage, and a flat bearing rail as an outrigger to provide moment stability. This is the classic *vee & flat* bearing that is still commonly used on many machine tools. However, this asymmetry also leads to asymmetry in the way the machine tool responds to the point of application of external loads, including the weight of a part. In the early 1900's, this gave rise to the use of the *double-vee* way; however, in order to get all four surfaces to actually contact, requires an exceptionally high level of worker skill to hand scrape the components, or exceptionally accurate grinding machines.

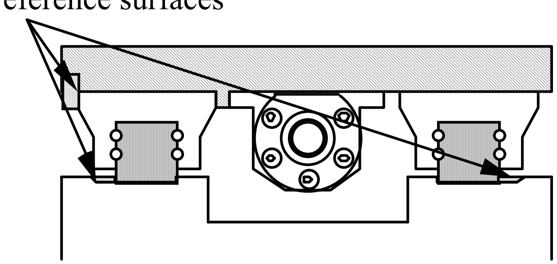
List what is symmetric and what is not symmetric about each of your potential ***concepts***, and ask yourself if *symmetry* helps or hurts. Now study what would happen if what was symmetric was made not symmetric, and vice versa by philosophically applying reversal.

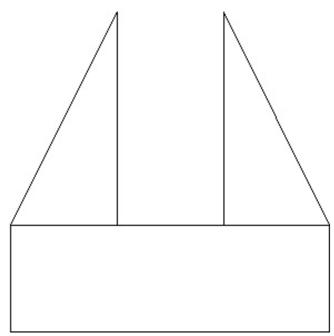


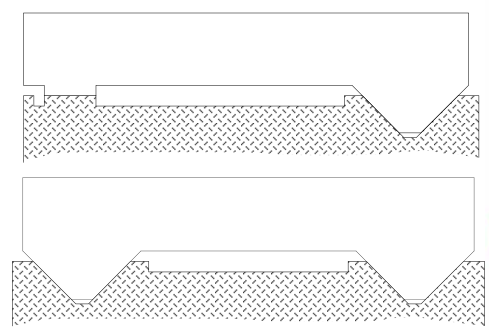
Symmetry

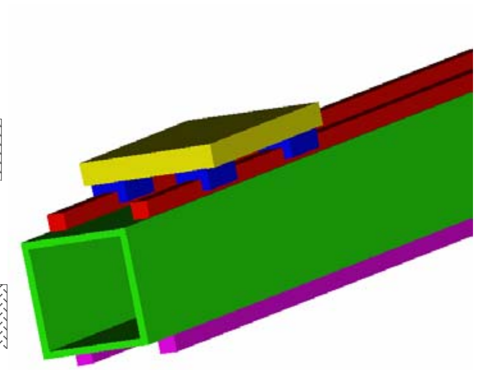
- *Symmetry* can be a powerful design tool to minimize errors
  - Thermal gradient errors caused by bi-material structures can minimize warping errors
    - Steel rails can be attached to an aluminum structure on the plane of the neutral axis
    - Steel rails on an aluminum structure can be balanced by steel bolted to the opposite side
  - Angular error motions can be reduced by symmetric support of elements
- *Symmetry* can be detrimental (Maxwell applied to symmetry)
  - Differential temperature minimized by adding a heat source can cause the entire structure to heat up
    - Only attempt with extreme care
    - Better to isolate the heat source, temperature control it, use thermal breaks, and insulate the structure
  - A long shaft axially restrained by bearings at both ends can buckle
  - Remember-when you generalize, you are often wrong
    - The question to ask, therefore, is "Can symmetry help or hurt this design?"

Blocks to push components against precision ground reference surfaces









Parallel Axis Theorem

The *Parallel Axis Theorem* is a very powerful tool for determining a complex structure's second moment of the area, commonly called the *moment of inertia* in the study of strength of materials, with respect to any axis. Starting with basic shapes, whose areas and moments of inertia about their local neutral axes are known, a design engineer first finds the location of the neutral axis, and then uses the parallel axis theorem to determine the moment of inertia for the entire section. Using a spreadsheet or MatLab, the design engineer can thus quickly optimize the design of a composite section or a truss. Composite sections (laminates) and trusses are of fundamental importance in the design of strong lightweight machines.

Consider sheet metal of thickness *t* bent into a channel of height *h* and width *w*. Where is the neutral axis and what is the moment of inertia? Start by finding the area of moments of inertia of simple rectangles that make up the channel (sketch this!) Pick a coordinate system coincident with the bottom of the beam and find the location of the neutral axis with respect to it. Then apply the parallel axis theorem:

$$I_{flange} = \frac{th^3}{12} \quad A_{flange} = ht \quad I_{web} = \frac{(w-2t)t^3}{12} \quad A_{web} = (w-2t)t$$

$$y_{NA} = \frac{\frac{h}{2}A_{flange} + \frac{t}{2}A_{web} + \frac{h}{2}A_{flange}}{A_{flange} + A_{web} + A_{flange}}$$

$$I = 2I_{flange} + I_{web} + 2\left(\frac{h}{2} - y_{NA}\right)^2 A_{flange} + \left(\frac{t}{2} - y_{NA}\right)^2 A_{web}$$

As a simple example, consider a laminate (see page 8-29) of metal on either side of a wood core, where the neutral axis is located at the center of the beam. A laminate is typically defined by its width *w*, the thickness *t* of the top and bottom chords (skin) and the thickness *h* of the spacing material (core), e.g., honeycomb, so the total thickness of the section is $H = h + 2t$. The core material's contribution to the stiffness is usually negligible but for completeness it is included at first:

$$I = \frac{wt^2 E_{chord}\left[\frac{t^2}{3} + (h+t)^2\right] + \frac{wh^4 E_{core}}{12}}{2tE_{chord} + hE_{core}} = \left.\frac{w\left[(h+2t)^3 - h^3\right]}{12}\right|_{E_{core} \Rightarrow 0} \qquad \frac{I}{c} = \frac{2I}{H}$$

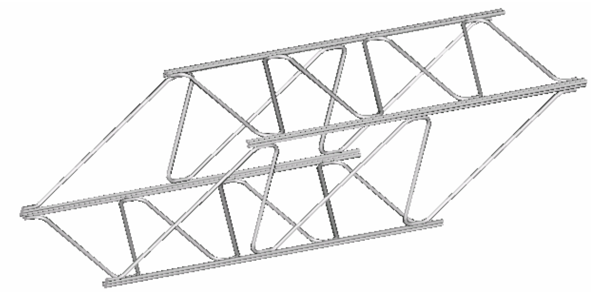
The parallel axis theorem enables you to quickly determine the maximum potential strength of a truss. The strength of a truss depends primarily on the top and bottom *chords*, and proper coupling between them using diagonal members called *braces*. The truss members that connect the chord members not only transfer shear between the chord members, they also control the free-length of the members in order to prevent buckling. For 2 round rods/chord, each of diameter *D*, and truss height *H*, the effective moment of inertia I_{truss} for determining the strength of the truss *(do not forget to also calculate the buckling loads in the chord members!)*:

$$y_{NA} = \frac{\sum_{i=1}^{4} y_i A_i}{\sum_{i=1}^{4} A_i} = \frac{2HA}{4A} = \frac{H}{2}$$

$$I_{truss} = \sum_{i=1}^{4} I_i + \left(y_{NA} - y_i\right)^2 A_i = 4\frac{\pi D^4}{64} + 4\left(\frac{H}{2}\right)^2 \frac{\pi D^2}{4} \approx \frac{\pi H^2 D^2}{4} \qquad \frac{I}{c} = \frac{\pi H D^2}{2}$$

These basic calculations allow a design engineer to rapidly determine if it is feasible to create a truss of an overall height and width that can fit into the design space and meet strength and stiffness goals. The design engineer can then move forward with the ***concept***, doing similar first-order calculations for other components. Once all such calculations are complete, and a design ***concept*** is found to be feasible, detailed design calculations can be done to determine truss member spacing and chord size to prevent buckling or determining laminate, core, and adhesive parameters to prevent delamination.

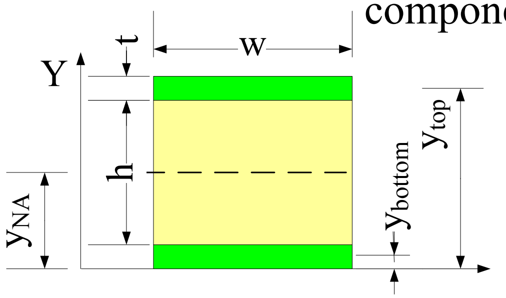
Identify high stress or deflection sensitive regions in your ***concept***, and do a first-order design for an appropriate laminate or truss. What is the difference in stiffness and strength between a pice of wood, and the same piece of wood with 1.5 mm thick steel or aluminum skin bonded to each side? What is the shear stress in an adhesive layer of a laminate you are considering? What might be the shear stress in the welds of a truss you are considering?



Parallel Axis Theorem

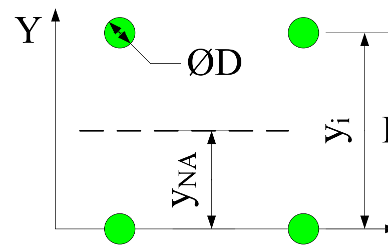
John McBean goes to the extreme!

- The *Parallel Axis Theorem* is useful for calculating the moments of inertia for complex objects
  - The stiffness of a design is proportional to the square of the distance of the component structural members' neutral axes from the assembly's neutral axis

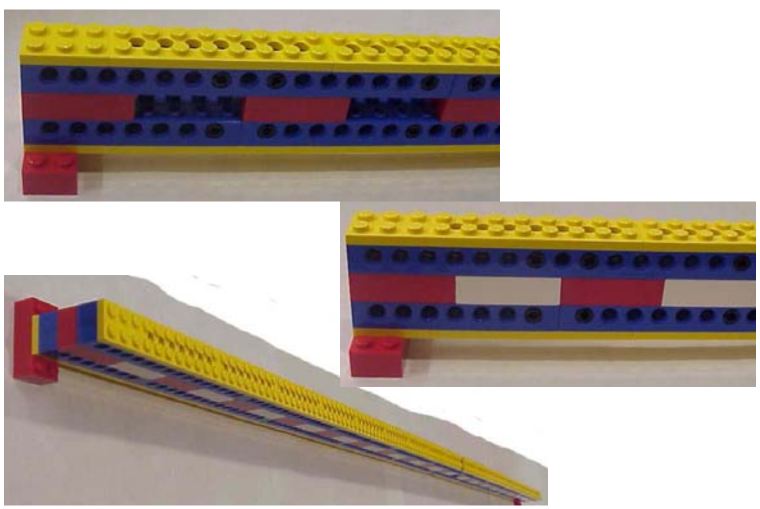


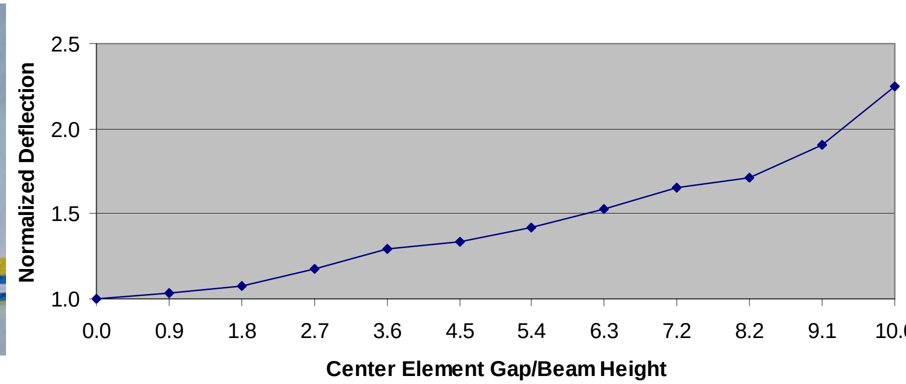
$$y_{NA} = \frac{\sum_{i=1}^{N} y_i A_i}{\sum_{i=1}^{N} A_i}$$

$$I = \sum_{i=1}^{N} I_i + \sum_{i=1}^{N} (y_i - y_{NA})^2 A_i$$



- The assembly's neutral axis is found in the same manner as the center of gravity, and it is located a distance y_{NA} from an arbitrary plane





Accuracy, Repeatability, & Resolution

The terms *accuracy, repeatability (precision)* and *resolution* are often improperly used, so it is important that their definitions be clearly understood. different types of machines require different types of performance, and since accuracy is generally more difficult to achieve than repeatability or resolution, large differences in design, and cost, can result.

Accuracy is the ability of a system to “tell the truth” or how well it can achieve or measure a state that is traceable to a national or international standard. For example, if one were to use a device to make multiple measurements of a known standard, the difference between the known measurement and the average of all the measurements taken would be the accuracy of the measurement. Note that the accuracy of measurements where the dominant error is random in general increases with the square root of the number of measurements taken; thus taking 100 readings and averaging them can increase the accuracy of the measurement by a factor of 10. Accuracy is affected by the shape and size of components, as well as thermal growth, backlash (gaps between elements), wear, friction, deformations, and sensor and control system accuracy.

Repeatability, or *precision,* is the ability of a system to “tell the same story every time” or how well it can achieve or measure the same state each time. For example, if one were to use a device to make multiple measurements of a known standard, the standard deviation of the measurements would be the repeatability of the measurement; and the number of standard deviations would give a confidence level about the repeatability. Repeatability is affected by thermal growth, backlash (gaps between elements), wear, friction, deformations, and sensor and control system repeatability.

Resolution, is the ability of a system to “tell the fine details of the story” or what is the minimum increment that can be detected or moved. For example, if one were to try and move to a desired position, and then measure the error, and then correct the position by the smallest achievable amount, it would be limited by the resolution of the system. Resolution is primarily affected by friction and also the resolution of the measurement and control system. On the super fine scale, surface roughness also becomes an issue.

In order to take full advantage of these definitions, always ask yourself when designing something “can the system be made with the desired accuracy?” What will be the required accuracy of each of the components, such as their size and shape and the squareness and straightness which they must be assembled? For example, engine components must bolt together, the crank bore machined, the components taken apart, and then other parts, such as the bearings and crankshaft, added and assembled to fit back together exactly, so that the two halves of the crank bore still form a circle. Since the engine parts are matched, that is the bedplate and the block are machined as a unit and then reassembled, repeatability is all that is required.

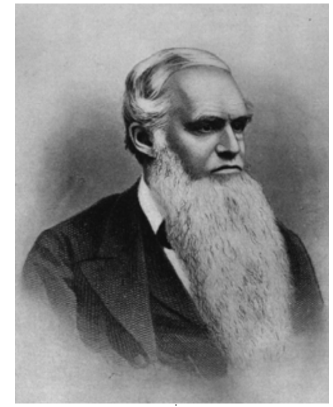
Accuracy, Repeatability, & Resolution are not just terms for mechanical performance, they are also philosophies for how to execute the design process. As noted by David Arguelles, the winner of MIT’s 1999 *MechEverest* design contest:

“My machine performed very well in the contest. It had zero malfunctions and we managed to hit dead on with the grapple hook every time. In round 4, we introduced the trailer addition to the machine. Everyone was surprised that someone could score more than 50 points. I got the highest score of the night at 58 points and went on to win the contest. Prof. Slocum seemed pretty excited that I had broken 50 points. Here he is picking me up and spinning me around in circles”.

Indeed, in competition after competition of just about any type, repeatability is the key to winning. The most clever design becomes a liability if it was finished the night before and has not been run dozens and dozens of times to find errors and maximize repeatability. It is the same principle with organizations. Training so critical functions are executed with near reflex-like action is the key to success.

Classify each of the motions your ***concepts*** require in terms of the required *accuracy, repeatability, & resolution*. How can the ***concepts*** be optimized to minimize required accuracy and instead rely more on repeatability? Can you minimize friction so as to likely increase resolution?

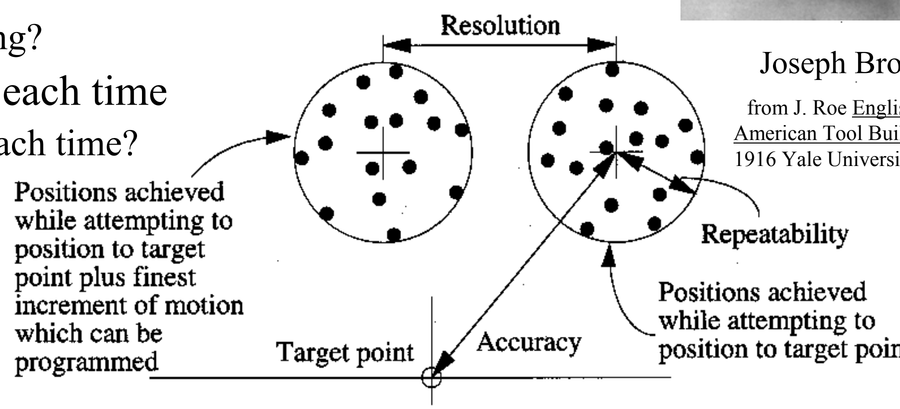
Accuracy, Repeatability, & Resolution



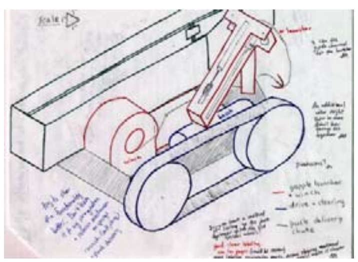
Joseph Brown

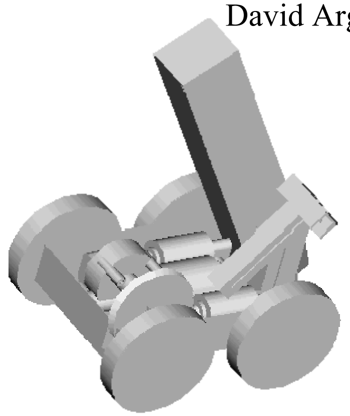
from J. Roe English and American Tool Builders, © 1916 Yale University Press

- Anything you design and manufacture is made from parts
  - Parts must have the desired accuracy, and their manufacture has to be repeatable
- *Accuracy:* the ability to tell the truth
  - Can two machines make exactly the same part?
  - Are the parts the exact size shown on the drawing?
- *Repeatability:* the ability to tell the same story each time
  - Can the machine make the exact same motion each time?
  - Are the parts all the same size?
- *Resolution:* the detail to which you tell a story
  - How fine can you adjust a machine?
  - How small a feature can you make?
- How do these affect the design process?

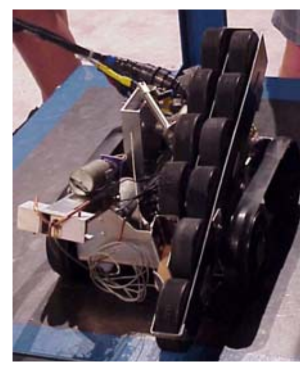


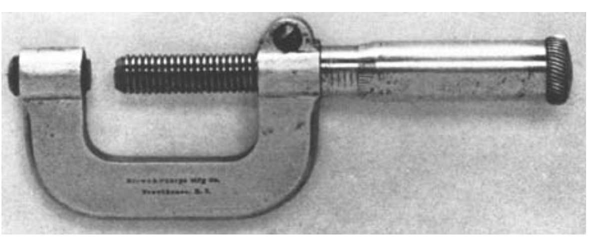
Hook launcher Model	
weight of hook (Kg)	0.05
muzzle velocity	9.4
Number of springs	2
d (draw)	0.095216
Winch model	
radius	0.05
mass	6
w (rpm)	55
torque	2.1
velocity	0.287833





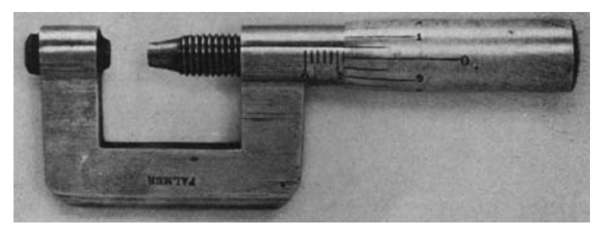
David Arguellis wins "MechEverest" with a machine that repeats every time!





One-inch Micrometer (left) made by Brown & Sharpe, 1868 and Palmer Micrometer (right) brought from Paris by Brown in 1867

from J. Roe English and American Tool Builders, © 1916 Yale University Press



Accuracy, Repeatability, & Resolution: *Mapping*

Mapping is the ability to measure or predict the repeatability or accuracy of a machine, and then use this knowledge of the error to compensate for it. Thus repeatability can sometimes substitute for accuracy. In the static sense, this means that if parts cannot be machined accurate enough to ensure that their assembly has the desired accuracy, then assemble them into the machine and then map the machine's performance. The machine's control system can use the map to compensate for the errors.

As introduced on page 2-6, consider the manufacturing of engines, where the bottom half of the engine, called the bedplate, is typically aligned to the block using dowel pins and then bolted in place prior to the crank bore being machined. Afterwards, the bedplate is removed and the bearings and crankshaft are installed, and the bedplate is repositioned and bolted to the block. If there is any non-repeatability in the bedplate-to-block alignment, the bearing journals will not be aligned, and engine quality will suffer. It takes 10 dowel pins to achieve 5 micrometer bedplate-to-block alignment repeatability. A quasi-kinematic coupling, however, uses three steel spherical button shaped elements pressed into the bedplate which then mate with three tapered holes with their sides partially removed to make them quasi vee-grooves. This design is deterministic, equations can be written to predict all the forces, and 0.5 micron repeatability bedplate-to-block alignment repeatability.

In the case of servo-controlled assemblies, since the 1970's, it has become more and more common to make machines mechanically repeatable, and then use their controllers to compensate for repeatable errors. These errors are measured as a function of machine position and operating load and mapped as a final step in the manufacturing process. Modern industrial robots are a prime example, where without error correction algorithms, they have accuracies on the order of 2-3 parts per thousand, e.g., a robot with a 2 meter reach may have 5 mm error. With error correction, a factor of improvement of 10-20X is achievable.

The magnitude of error correction depends on the types of errors and how they are generated. In systems with mechanical contact between the elements, such as sliding or rolling element bearings, mapping can generally increase effective accuracy by a factor of 5-20. The greater the friction in the system, or the more heavily it is loaded, the lower the repeatability. However, it is interesting to note that if a system is too lightly loaded, it will also have poor repeatability, because conditions must exist where the largest number of elements remain just equally loaded. Consequently, heavily loaded machines tools may only achieve a factor of 5 improvement in accuracy by mapping, while modestly loaded machines may achieve a factor of 10 improvement. Lightly loaded machines may achieve a factor of 20 improvement.

Machines that use non-contact type bearings, such as aerostatic (air bearings), hydrostatic, or magnetic bearings, more typically achieve a factor of 20-50X improvement in accuracy, where the factor depends on the temperature control of the system, and the degrees of pressure or electronic control of the bearings. The figure shows a rotary hydrostatic bearing table, where the mechanical elements are only round to 2 microns, but the externally pressurized oil film that keeps the components separate from each other, allows sub-micron accuracy and tens of nanometer repeatability to be achieved. Flexural bearings also have an extremely high potential for having their errors mapped.

As introduced on page 1-15, the linear motion system, called an *Axtrusion*, is shown here with a map of its pitch motion. The pitch motion has a spatial frequency equal to that of the linear electric motor's permanent magnets; thus, even though the pitch error is on the order of 2 arc-seconds (9 microradians, or a slope of a 9 millionths of a meter per meter), its pitch repeatability is on the order of 0.2 arc-seconds. A machine tool would have, for example, three orthogonal axes, so the error in one axis, could be compensated for with motions in the other axes. An analytical model of the machine would be made using homogeneous transformation matrices, as discussed on page 3-10, and incorporated into the machine tool controller to allow the controller to calculate sine errors and then direct an axis to compensate for the error. For example, if the pitch error was not compensated for by the Axtrusion axis, at a point 0.3m from the bearing surface, a sine error along the axis of motion of 0.3m x 9 microradians = 2.7 micrometers. If the pitch error as a function of position along the axis had been mapped, then the axis itself could be caused to adjust its position to compensate for all but 0.3m x 0.9 microradians = 0.3 μm.

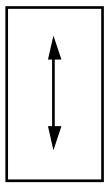
How might your ***concepts*** be made more repeatable, and how could this be used instead of requiring accuracy? What motion resolution do your ***concepts*** require and how might you reduce friction to help increase resolution potential?

Accuracy, Repeatability, & Resolution: *Mapping*

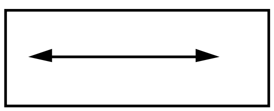
- It is often most important to obtain mechanical *repeatability*, because *accuracy* can often be obtained by the sensor and control system
  - When the error motions of a machine are *mapped*, the controller multiplies the part height by the axis' pitch & roll to yield the sine error for which orthogonal axes must compensate

Eli Whitney

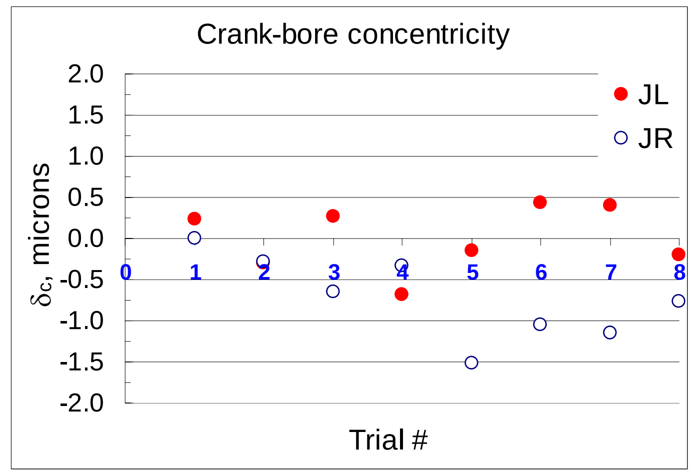
from J. Roe English and American Tool Builders, © 1916 Yale University Press

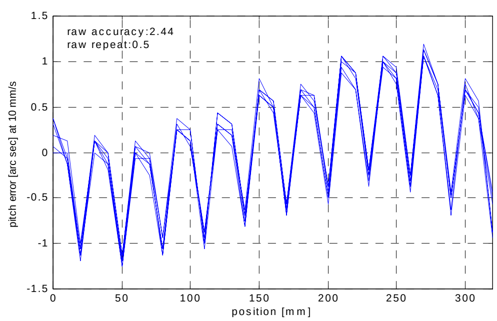


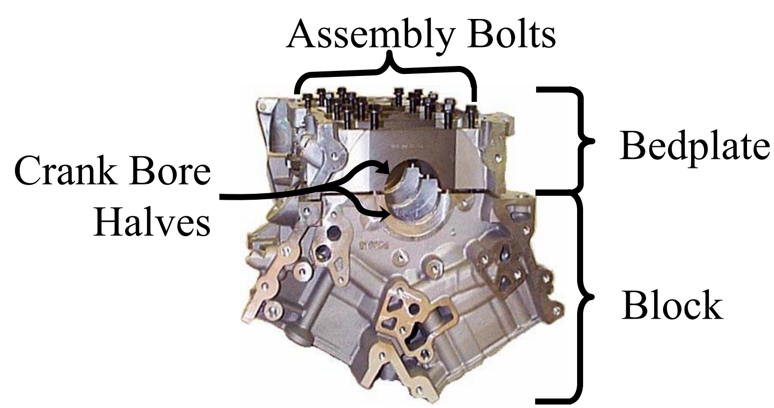
Y axis: Can be used to compensate for straightness errors in the X axis.



X axis: Can be used to compensate for straightness errors in the Y axis.







Sensitive Directions & Reference Features

A precision machine made up of many parts represents a host of challenges (opportunities!); therefore, in the spirit of Occam's razor, one should minimize the amount of effort required to achieve the design by determining where to focus design effort. Accordingly, one should always identify the directions in which accuracy or repeatability are most important. These are the *sensitive directions* in a machine. In addition to the sensitive directions, there are also key features, which will not change even though many other features in the design may change as it evolves. These *reference features* can be real, such as an edge, or virtual, such as a datum plane on a drawing.

Consider a tool, which is moved radially (X-axis) and axially (the Z-axis is aligned along the spindle's axis of rotation), to machine a part in a lathe. As the axes move, any errors in the axes motion along the radial or axial direction will directly affect the radial or axial accuracy of the part being machined. The radial (X) and axial (Z) directions are thus the sensitive directions of the machine and particular care must be taken to minimize errors in these directions. This means that the X-axis' actuator and sensor must be made very good, and the Z-direction straightness of the X-axis bearings must also be made very good. Similarly, the Z-axis actuator and sensor must be made very good as well as the Z-axis' X-direction straightness! In both cases, however, errors in the Y direction do not directly cause radial or axial errors in the part, and thus the Y direction is a non-sensitive direction.

Even though sensitive directions highlight where particular attention should be paid when designing the system, one does not blatantly ignore the other directions. On the contrary, this just means that one applies higher quality elements, such as bearings, actuators or sensors, to the sensitive directions and one spends less money on the non-sensitive directions. This naturally leads to the concept of *reference features* in a physical design, and in a CAD model. Robust designs have physical reference features from which measurements are made with respect to, and to which components are attached. Consider the case of holes in a part, where the designer must first determine the *design intent* of the holes. If the holes are part of a large series of bolt holes, then the position of each hole with respect to a reference edge (feature) is required so they will line up with the holes in another part. On the other hand, if the holes are to receive precision parts that align them, such as dowel pins which are press-fit into the holes, it is most important to minimize the possible error between the holes, and one hole's position should be referenced with respect to the other hole.

Consider typical machining tolerances of 0.1 mm for any single linear dimension. In the former case, this means any single hole may be improperly positioned by 0.1 mm. However, this also means that there is a potential for the distance between any two holes to be in error by 0.2 mm, if one hole is on the minus side of tolerance and the other is on the plus side of the tolerance. On the other hand, if one hole is dimensioned with respect to the other, then the error in position between them will be at most 0.1 mm. This is why it is so important when creating solid models in the concept phase to always consider the *design intent* of every feature, and dimension each feature to capture the design intent. All too often, design engineers quickly choose the dimension that is easiest to click on so they can rush through the design ***concept*** phase. However, it is during the ***concept*** phase that careful thought should be paid to capturing the *design intent*, so when it comes time to do the detailed design, it can be done with accuracy and ease.

Anybody can create a solid model with lots of colorful elements, but will the solid model still work if features or components are deleted? Do the dimensions capture the true intent of the design to achieve the accuracy required? Can properly dimensioned drawings be automatically rapidly created because each of the features was defined using dimensions made with respect to reference features in the first place? The wise manager will review how an employee creates a solid model, and not just look at the pretty picture that is shown at an interview. Reciprocity also shows that the wise job candidate will print a picture of a solid model with the dimensions shown so they can show that they know how to build a robust solid model and create robust designs! Remember, "Random results are the result of random procedures"[1]

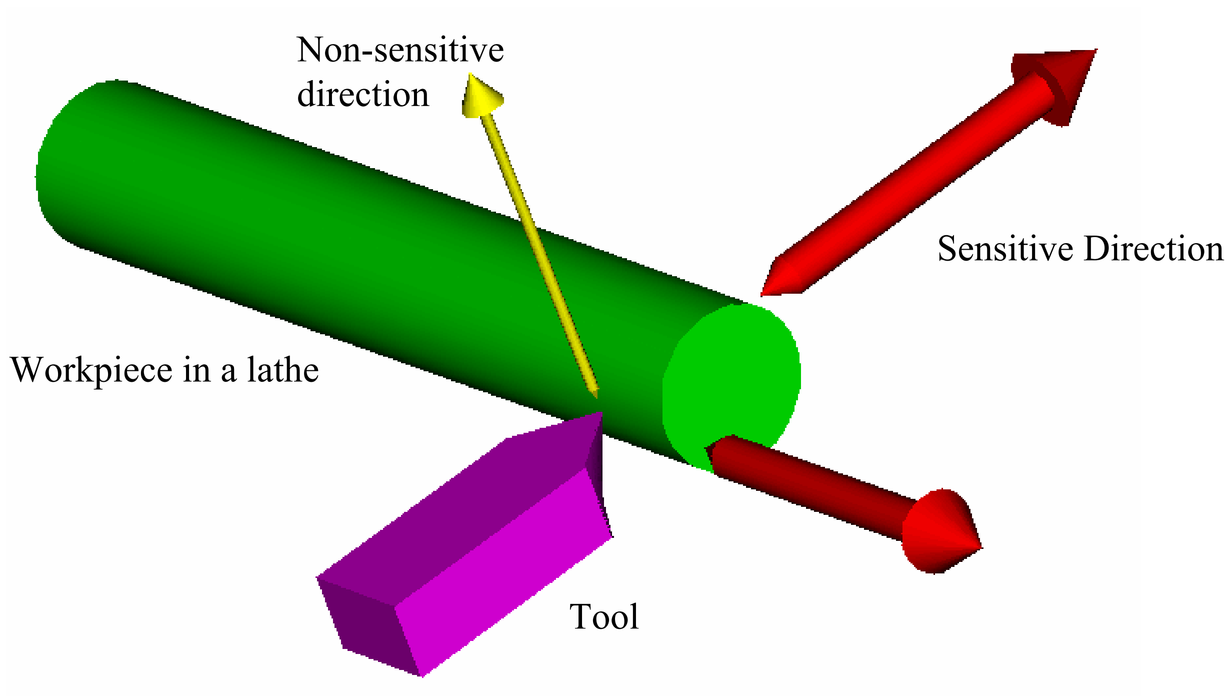
What are the sensitive directions for your ***concept***, and therefore what elements of the design must be most carefully engineered and manufactured? What reference features can you use to properly and robustly dimension features of your design to ensure you realize each features' design intent? What reference features might you use to ensure proper dimensioning of the part drawings when they are later generated from the solid model?

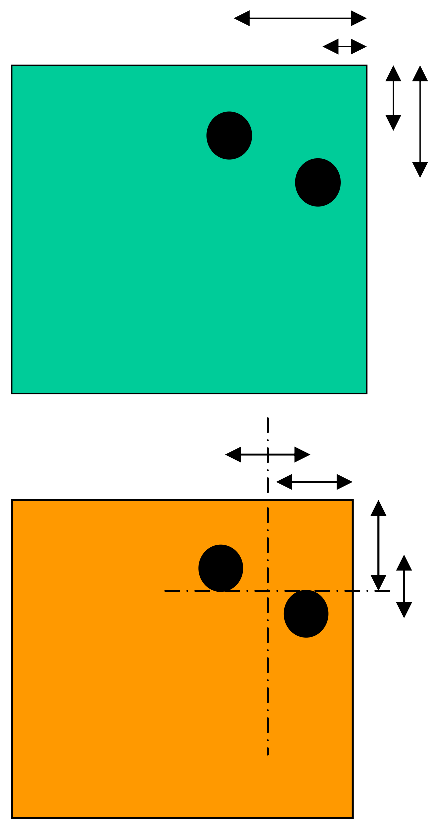
---

1. Geoffe Portes, Cranfield Unit for Precision Engineering (CUPE), Cranfield, UK

Sensitive Directions & Reference Features

- In addition to *accuracy*, *repeatability*, and *resolution*, we have to ask ourselves, "when is an error really important anyway?"
  - Put a lot of effort into accuracy for the directions in which you need it
    - The *Sensitive Directions*
    - Always be careful to think about where you need precision!





Structural Loops

A machine is nothing more than an assembly of parts to guide the motion of one object with respect to another, and in so doing, to provide a means to support forces and moments according to Newton's laws. Accordingly, the path that forces take through the machine's structure to connect the tool to the work piece is called the *structural loop*[1]. One can tell a great deal about a machine merely by looking at its *structural loop*. The first and foremost observation to be made is that the smaller the structural loop, the more likely the machine will have a high degree of stiffness. Secondly, if one merely draws a line on the sketch of a machine that traces the path the forces flow, one can gain an idea of machine performance. Given the structural loop for the machine in different configurations, the less the length of the structural loop varies, the less the machine stiffness will vary during operation, and the more likely it will achieve consistent performance.

Consider the two machine tool concepts shown, where one is designed for constant stiffness and one for maximum stiffness. The machines appear very similar, but if one looks closely, one sees that the Z-axis, which in a machine tool is always aligned along the spindle's axis of rotation, that supports the spindle has its motion guided in one of two ways. In the constant stiffness design, the bearing blocks are attached to the Z-axis carriage, and they ride on the bearing rails fixed to the Y-axis carriage. In the maximum stiffness design, the bearing blocks are fixed to the Y-axis carriage, and the bearing rails are attached to the Z-axis carriage.

In the maximum stiffness design, when the Z-axis is fully retracted, the tool is close to the bearing blocks and the structural loop shows a short path for the forces to flow through the machine to connect the tool and the workpiece. In this position, the machine has a high degree of stiffness, and a machinist could use the machine to achieve the highest of accuracies, because machining forces would cause minimal deflections in the machine. On the other hand, if the machine where being used to machine a large part, as the Z-axis extended, its structural loop is seen to lengthen considerably, and its stiffness would decrease. Thus under the same machining conditions, the error would increases, resulting in taper in the part. A machine could compensate for this by making multiple passes so the last cut was made with minimal force; however, this would not be suitable for production part manufacture. The maximum stiffness design is thus most suited for a machine where the design intent is to provide a machinist with the ability to make very precise difficult-to-machine parts.

In the constant stiffness design the length of the structural loop is nearly equal in both the extended and retraced positions. These longer, but constant length, structural loops result in a machine with less stiffness, but this also means that the deflection of the machine is more likely to be constant across its range of motion, and hence a large part will be made with less angular error (less taper). Less taper means less amplification of angular errors. Thus this type of machine is most suited for production use.

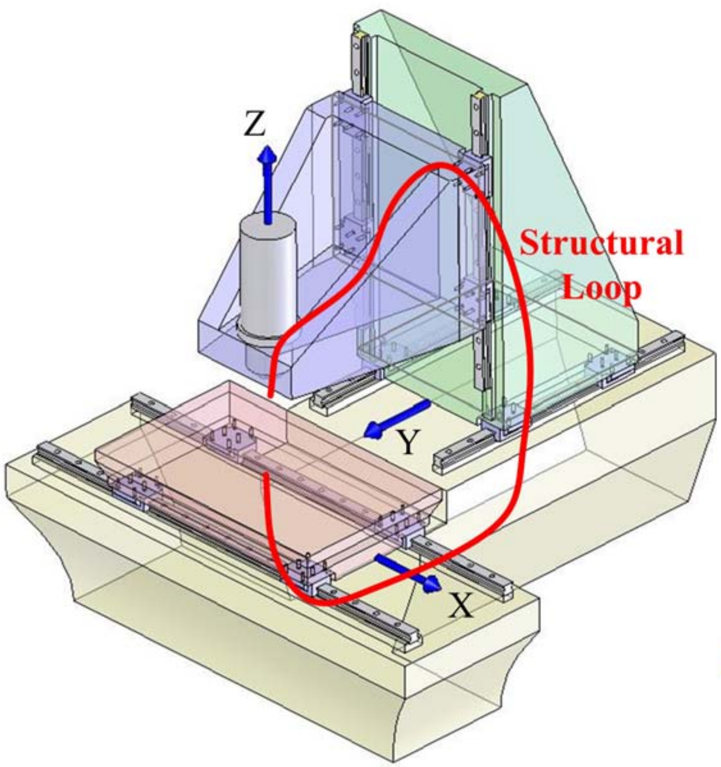
One can also use the length of the structural loop and the anticipated quality of components and manufacturing capabilities to rapidly estimate the accuracy potential of a machine. Accuracy of components and manufacturing processes is typically represented in parts-per-million. A typical milling machine in a student machine shop can make parts that have an accuracy of one part per thousand, which means a 100 mm part will have features on it accurate to 0.1 mm. If a collection of these parts is machined and assembled, and the structural loop is 400 mm long, then one might expect the accuracy that the machine might achieve would be on the order of 0.4 mm. This assumes of course that attention is paid to the detail of elements such as bearings and actuators. The inverse is also true, where if you have an estimate for the structural loop length, and you know the accuracy you are trying to achieve, you can estimate the quality level of components and manufacturing processes you will need. 1000 ppm is generally easy to achieve. 100 ppm requires careful attention to fixturing and machining process, but can be achieved by a skilled machinist on good quality machines. 10 ppm requires a highly skilled machinist working on very high quality machines, generally in temperature controlled rooms. 1 ppm requires highly specialized machines, machinists and facilities. In most cases, if the parts can at least be measured, then hand finishing techniques can be used to achieve order of magnitude increases.

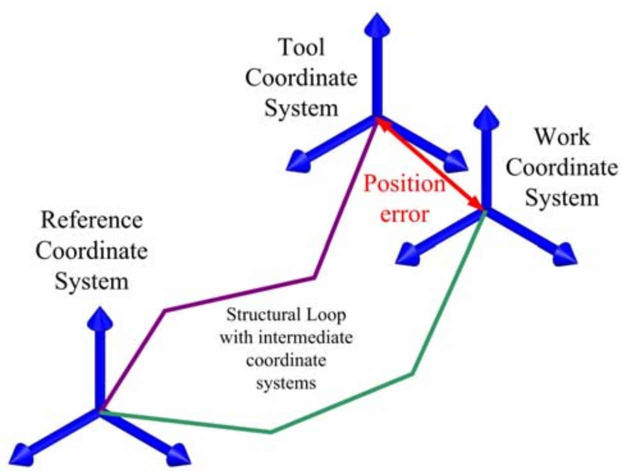
Trace the path of the flow of forces through each of your different ***concepts*** in each of their different configurations. How do the structural loop lengths differ and what might this indicate about machine performance? What machine accuracy is required and therefore what do different ***concepts***' structural loop lengths imply about respective required manufacturing accuracies?

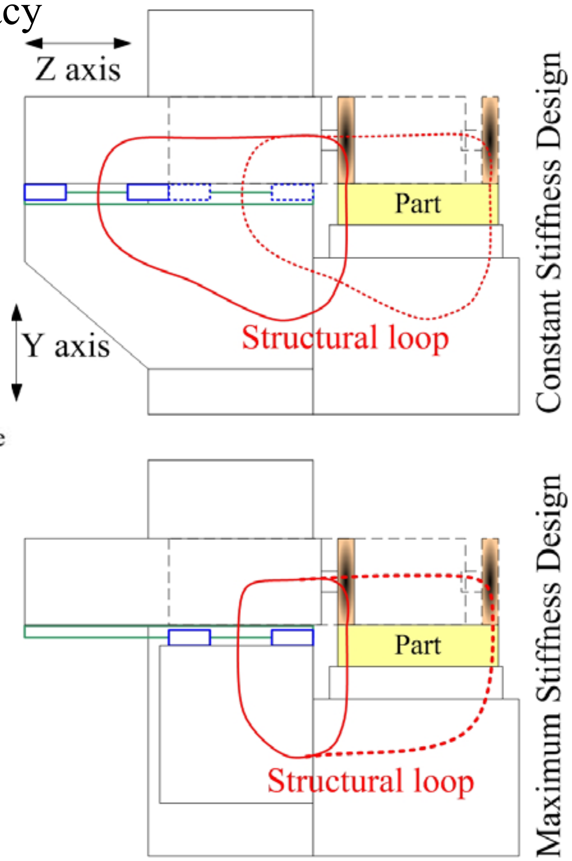
1. This would also make a good name for a breakfast cereal or a rock & roll band!

Structural Loops

- The *Structural Loop* is the path that a load takes from the tool to the work
  - It contains joints and structural elements that locate the tool with respect to the workpiece
  - It can be represented as a stick-figure to enable a design engineer to create a ***concept***
  - Subtle differences can have a HUGE effect on the performance of a machine
  - The *structural loop* gives an indication of machine stiffness and accuracy
    - *The product of the length of the structural loop and the characteristic manufacturing and component accuracy (e.g., parts per million) is indicative of machine accuracy (ppm)*
    - Long-open *structural loops* have less stiffness and less accuracy







Preload

Components which move relative to each other do so by means of bearings, and clearance typically exists between bearing elements to enable motion to occur. The need for clearance exists, for example, because a shaft may not be perfectly round. In addition, its shape may change slightly due to external loads or thermal expansion. Similar factors can affect the size of the hole (bore) in a bearing. Thus the bore must be larger than the shaft diameter in order to accommodate the worst-case scenario. However, this creates the condition whereby the exact radial position of the shaft's axis of rotation is not known. For many components, such as robot wheels, this may not pose a problem. On the other hand, for gears, this may prevent the teeth from properly meshing. In addition to radial error motions described above, axial clearance, and axial error motions, also exists.

A similar condition exists for linear motion systems, where the bearing rails, used to guide the motion and on which the bearings ride, may not be of uniform size and shape along their length, nor may they be exactly parallel. As a result, clearance typically exists between the bearings and the bearing rails to accommodate the maximum possible deviations. However, this also results in potentially unwanted error motions between the carriage supported by the bearings and the bearing rails.

And the bad news does not stop here, for these types of unwanted error motions caused by clearance between components, often referred to as *backlash*, occur between gears, leadscrews and nuts, and many other types of components. A common backlash problem is in geartrains which are used, for example, in steering systems. Anyone who has ever driven a very old car or a beat-up go-kart has experienced the fun of steering the vehicle when you turn the wheel some and nothing happens!

Fortunately their are two strategies that can be used to limit or prevent backlash: a) increase the accuracy of components and decrease clearances; or b) preload components against each other. The former can also increase the system accuracy (see page 3-17), but increases cost and can decrease system robustness in the presence of large deformations and differential thermal expansion of components. The latter is generally not that difficult to accomplish, greatly increases system reliability, and can be a very robust solution.

A typical method for preloading components is to select one component as the rigid element, and the other as the compliant element. The rigid element is generally selected to the element that will bear the greatest load in the system. The amount of preload provided by the compliant element should typically be just greater than the load that will be applied to unload the preload. If the applied load is less than the preload, then the system will act as a body supported by two springs, and the free-body-diagram yields the following equilibrium condition:

$$F_{load} - \left(F_{preload} + k_{\text{stiff element}}\delta\right) + \left(F_{preload} - k_{\text{compliant element}}\delta\right) = 0$$

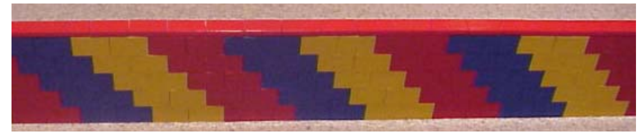
From this and the relation $F_{load} = K_{total}\delta$ the total stiffness of the system is shown to be equal to the sum of the rigid and compliant members, and this is true in either direction of loading, so long as in the direction that unloads the preload, the applied load is less than the preload force:

$$k_{\text{total}} = k_{\text{stiff element}} + k_{\text{compliant element}}$$

This basic philosophy can be applied to any mechanical system, and in many different ways. The challenge to the design engineer is to develop clever methods of creating compliant member which deform when subject to local distortions Δ while not introducing undesirable deflections in other directions. Furthermore, to reduce the variation in the preload force on the system, the amount of deflection used to create the preload by compressing the compliant element should be substantially larger than any expected local deformations.

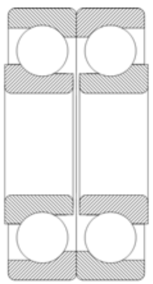
Where might backlash be a cause for concern in your robot? Might it cause you to over compensate when you give a control command and the system does not respond, so you give a larger command and all of a sudden too much motion occurs? Which ***strategy*** would be least costly for you to implement? Can simple extension or compression springs be used to preload critical ***components***? Is there any danger of over constraining your system by preloading it, such that in some conditions it may lock up and cease to move? Could preloading the system result in high joint friction forces that exceed the limits of your actuators, or cause them to use up too much power?

Preload

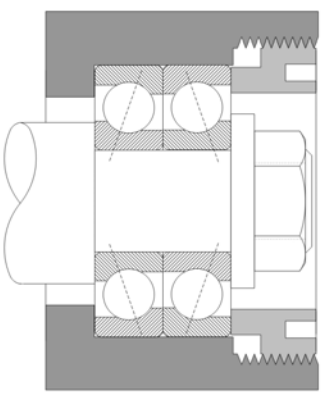


- Components that move relative to one another generally have tolerances that leave clearances between their mating features
  - These clearances result in *backlash* or wobble which is difficult to control
    - An example is the Lego roller coaster on page 3-10

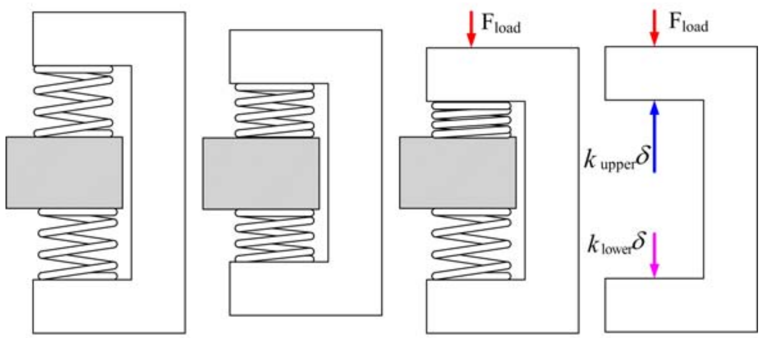
- Because machine elements often have such small compliance, and to account for wear, backlash is often removed with the use of *preload*
  - Preload involves using a spring, or compliance in the mechanism itself, to force components together so there is no clearance between elements
    - However, the compliance in the preload method itself must be chosen such that it locally can deform to accommodate component errors without causing large increases in the forces between components
      - Linear and rotary bearings, gears, leadscrews, and ballscrews are often preloaded
        - » One must be careful when preloading to not too over constrain the system!
      - Structural joints are also often preloaded by bolts



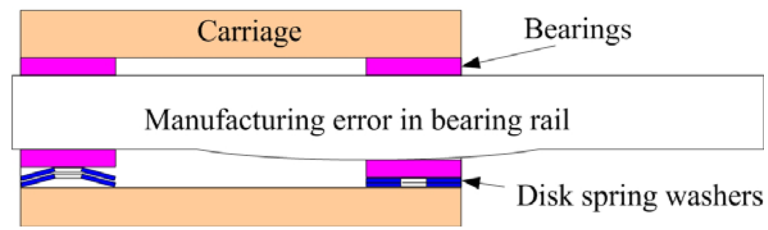
Bearings before mounting (inner ring axial clearance exaggerated)



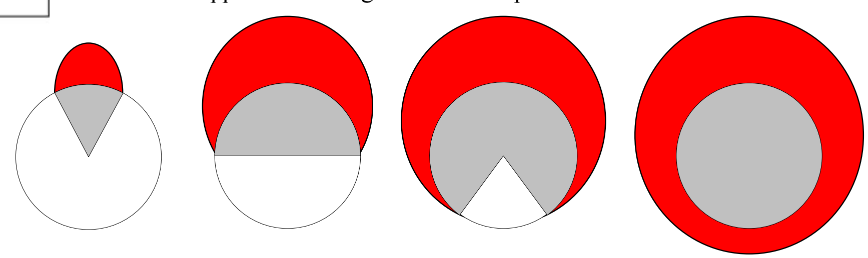
Back-to-back mounting after inner rings are clamped together



Bearings just touching rail | Bearings preloaded | Carriage fully loaded



Load distribution on rolling elements due to radial load applied to bearings with various preload conditions



Internal clearance | Zero clearance | Light preload | Medium preload

Centers-of-Action

Centers-of-action are virtual points within a body where various physical parameters can be modeled as lumped parameters such that forces applied through these points generate no moments on the body. Conversely, these are also the points about which angular motion occurs when forces are applied elsewhere on the body. Minimizing moments on a system minimizes angular motion, which minimizes sine errors and thus has the potential to enhance the robustness of a design. These centers-of-action include the body's mass as well as the stiffness and friction properties of its interfaces to other bodies. In addition, the thermal growth of a body can be analyzed with respect to a point which appears to remain fixed. The center-of-action of a group of masses (m_i), or a group of springs (k_i) or a group of friction forces (F_f) is found with respect to any arbitrary coordinate system by:

$$X_{coa} = \frac{\sum_{i=1}^{N} \zeta_i x_i}{\sum_{i=1}^{N} \zeta_i} \qquad Y_{coa} = \frac{\sum_{i=1}^{N} \zeta_i y_i}{\sum_{i=1}^{N} \zeta_i} \qquad Z_{coa} = \frac{\sum_{i=1}^{N} \zeta_i z_i}{\sum_{i=1}^{N} \zeta_i}$$

The *center-of-mass* is the point where an applied force only causes linear acceleration. No angular acceleration occurs which would lead to sine errors! The center-of-mass of a system of particles moves like a single particle of mass $M = \Sigma m_i$ under the influence of the resultant external force acting on the system. When a vehicle is loaded, if the center-of-mass is between say the wheels and the donkey's hooves, then the cart is stable. However, if more and more stuff is piled onto the back of the cart, the center-of-mass shifts to the rear of the cart. When the center-of-mass is over the wheels, the cart is neutrally stable. One more straw and the donkey goes up as the system becomes unstable! Similarly, when a vehicle drives up an incline, see the example on page 3-7, it will not tip over if the downward projection of the center-of-mass remains within the wheelbase.

In a similar manner, a body supported by bearings of stiffness k, behaves as if all the bearings' stiffnesses, both linear and the effective moment stiffness created by bearings spaced apart, were concentrated at the system's *center-of-stiffness*. When a force is applied to the center-of-stiffness, no angular motion of the structure occurs, thereby preventing sine errors from cascading through the system. Thus locating bearings in the plane of applied forces in a machine can help to minimize angular deflections and resulting sine errors. In a reciprocal manner, if a machine element (e.g., a leadscrew nut) is attached at the center-of-stiffness, then error motions of one machine element (wobble of the screw) will not cause pitch errors (sine errors) in another element (carriage actuated by the leadscrew). As an example, consider two springs k_1 and k_2 spaced a distance L apart. The center-of-stiffness is located a distance a from spring k_1. The location of the center-of-stiffness from spring k_1 and the equivalent lumped linear and angular stiffnesses at this point are found from a simple force and moment balance:

$$a = \frac{Lk_2}{k_1 + k_2} \qquad k_F = k_1 + k_2 \qquad k_M = \frac{k_1 k_2 L^2}{k_1 + k_2} \qquad K_F = \sum_{i=1}^{N} k_i \qquad K_M = \sum_{i=1}^{N} a_i^2 k_i$$

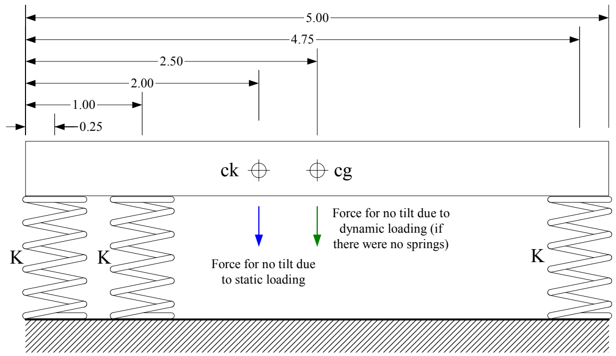
Similarly, the *center-of-friction* is the point where a force applied to a structure supported by bearings causes no angular motion of the structure. It is also found using force and moment balance equations that consider the effects of friction, bearing geometry, and center of gravity. The center of friction is sometimes, but not always, located at the center-of-stiffness. If a load is applied to different positions on a Vee-and-Flat and a Double-Vee supported carriage, how do the center-of-friction and the center-of-stiffness vary? How do they compare to a box way design? What are the cost/benefits of the two designs?

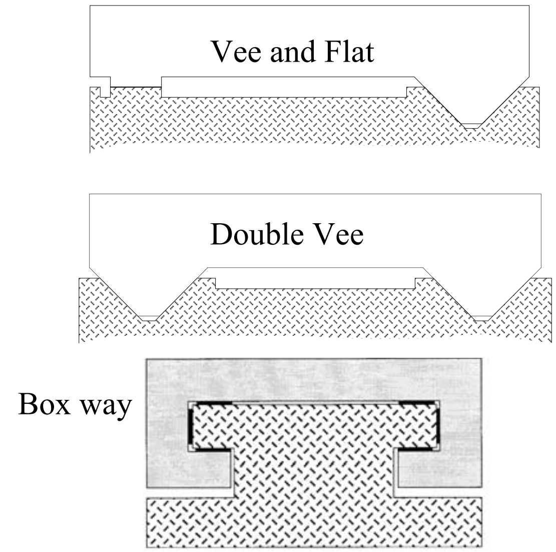
If the machine's actuators and work zone are near the machine's components' centers of mass, stiffness, and friction, errors can be reduced. Even though it may not be possible to always totally achieve these goals, getting as close as you can will make your job of completing the detailed design of the machine a lot easier, and the machine is much more likely to be robust. Thus the principle of centers-of-action allows a design engineer to create a ***concept*** as a stick figure to be filled out later. This helps prevent going into too much detail too early on in the ***concept*** phase. It also allows a design engineer to simplify the model of a complex assembly.

To help hone your skills, derive the above formula for the lumped angular stiffness for the two spring system and the generalized formula for N springs supporting a rigid body. Trace the *structural loop* and try to envision the dominant sources of compliance (lowest stiffness elements) in each of your different ***concepts***. How might angular deflections cause a decrease in machine performance?

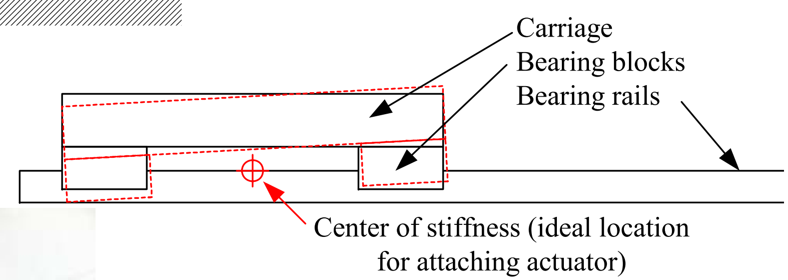
Centers-of-Action

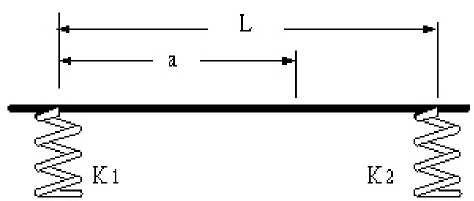
- The *Centers-of-Action* are points at which when a force is applied, no moments are created:
  - *Center-of-Mass*
  - *Center-of-Stiffness*
  - *Center-of-Friction*
  - *Center-of-Thermal Expansion*

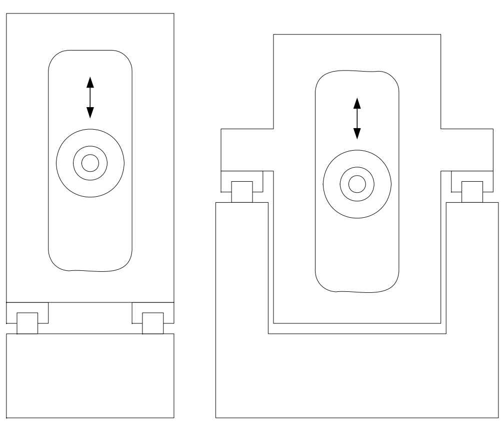




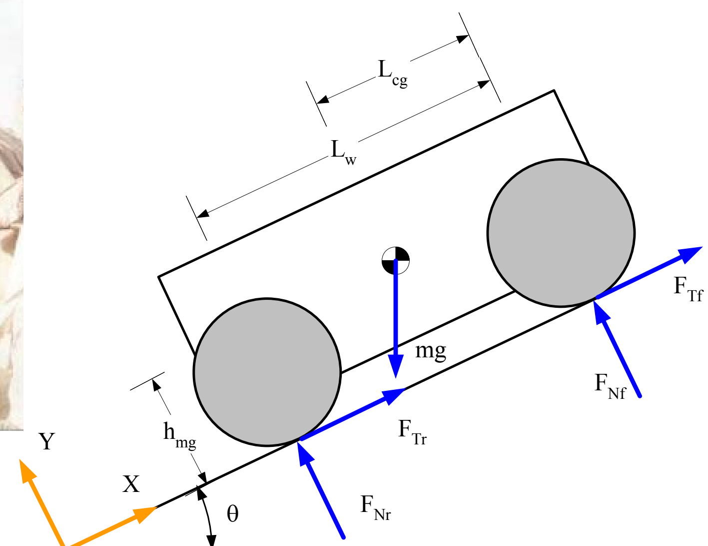
- A system is most robust when forces are applied as near as possible to the *Centers-of-Action*







Funny image found on http://zeeb.at/oops/Nothing_Changes.jpg, photographer not credited, would like to, email slocum@mit.edu



Exact Constraint Design

Much of engineering is based on being able to analyze a design before it is built. Thus is it not better to design something that can be analyzed?. *Exact Constraint Design* is based on the principle, that every design's structure should be statically determinate. When a component is supported in a statically determinate manner, it is *kinematically located,* and the design is *kinematic*.

Since six unknowns can be obtained from the solution of six independent simultaneous equations (ΣF_x, ΣF_y, ΣF_z, ΣM_x, ΣM_y, ΣM_z,), the reactions of a stable mechanism having exactly six unknown reaction elements may be obtained from the simultaneous solution of the six equations of static equilibrium. In such a case, the reactions of the structure are *statically determinate.* A design that is statically determinate is said to be an *Exact Constraint Design,* and the chances are good that the design can be engineered to be very robust. The downside is that such designs often cannot take advantage of symmetry, and the resulting asymmetric loading may produce asymmetric deflections that prevent the machine from reaching its desired performance level. If the reaction elements, forces and moments, cannot be determined from six equilibrium equations, then deformations need to be considered, and the structure is said to be *statically indeterminate*. A design that is statically indeterminate is an *Over Constrained Design,* and there is a *chance*, but not always, that the design might not be very robust. The upside is that such designs can often have overall greater strength and stability, and can take advantage of symmetry.

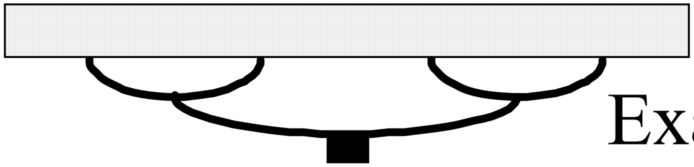
A prime example is that of a chair. Thousands of years ago, three legged stools evolved. Perhaps this occurred because rocking on four legs created such an annoying sound that Og broke one of the legs off of his stool to hit Zog on the head; hence was born the club and the deterministic stool? Since a four legged chair or table often has a rocking motion, why would we then would we want a five legged chair? In fact, a common design challenge that manifests itself when ***components*** are assembled, is that as fasteners are engaged, elements can deform, and bearings can become overloaded. This is because designers do not always carefully consider that a machine is like a collection of springs, and if you push on one spring, the entire web will deform in order to come to equilibrium. A system may even be assembled so carefully that the tightening of bolts, for example, does not cause any overloading. But then the system temperature changes and thermal expansion can cause bearings to be overloaded (see page 10- 27) or shafts to buckle.

So what is a design engineer to do? The answer is to identify sensitive directions in the system, and then pick an anchor point to minimize system errors, and then pick another point to constrain rigid body motion, yet allow for expansion or misalignment. Consider a shaft to be supported by bearings. The shaft near the most sensitive element to which it connects should be supported by bearings that resist both radial and axial loads. The other end of the shaft, about 3-5 diameters away according to Saint-Venant, should be supported by bearings that only provide radial load support. Ideally, both sets of bearings would allow for slope errors, as discussed in the example on page 3-4. If these bearings also provided axial load support, deformations in the support structure would likely axially load the bearings and cause them to fail. An exception would be where one bearing in a back-to-back set is placed at each end of the shaft. In high speed systems, the shaft generally gets hotter than the outer structure and thermal expansion of the shaft can cause the bearings to fail.

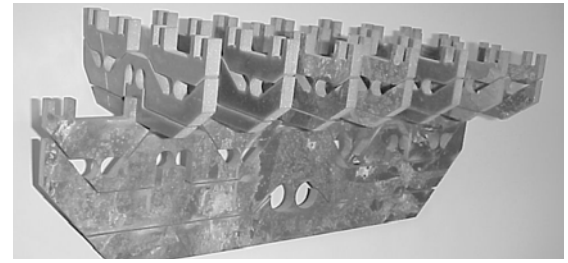
Is this always the case? In many advanced machine systems, long shafts are held in tension by bearings that resist axial forces at both ends of the shaft. The shaft is tensioned such that mechanical strains offset any thermal strains, and the tensioned shaft has higher natural frequency so it can rotate faster and be axially stiffer. This is an advanced design concept that has its own very special set of design principles (more to learn in your next class!) Can a design be both exactly constrained and elastically averaged? A windshield wiper's main functional requirement is to provide uniform loading on the wiper blade as it traverses a surface with varying curvature. The design parameter that make this possible is the *wiffle tree* which takes a single point force and spreads it out via a diverging mechanism[1]. In three dimensions, a large heavy precision surface plate can be supported at three points, so it will be as flat as when it was made, and this is made possible by using a wiffle tree!

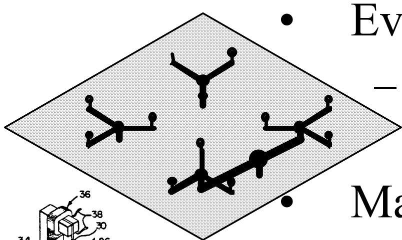
Study the photos of the motor with mounts at the front and at the rear. if the Motor is mounted to a plate that is not flat, which mount will not cause the gearbox to deform? If the gearbox deforms, what happens to quality?

1. Also spelled whiffle tree. See for example http://www.foothill.net/~sayre/22-in.%20binocular%203.htm

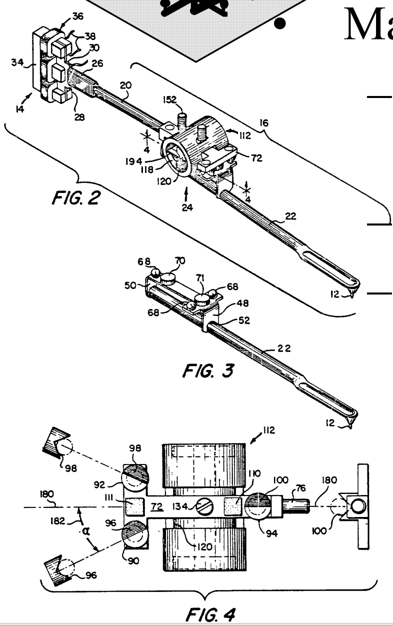


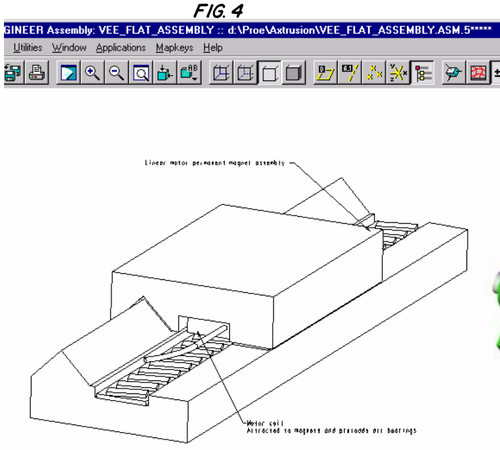
Exact Constraint Design

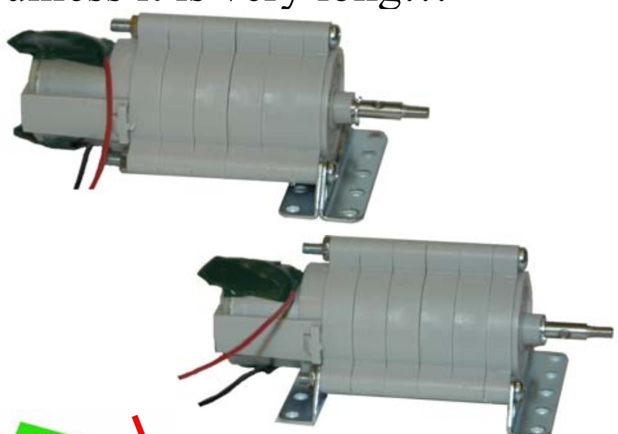


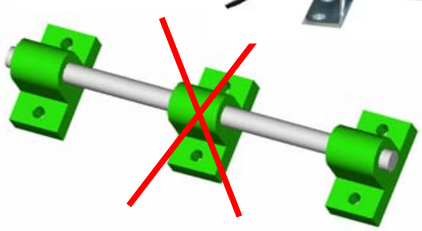


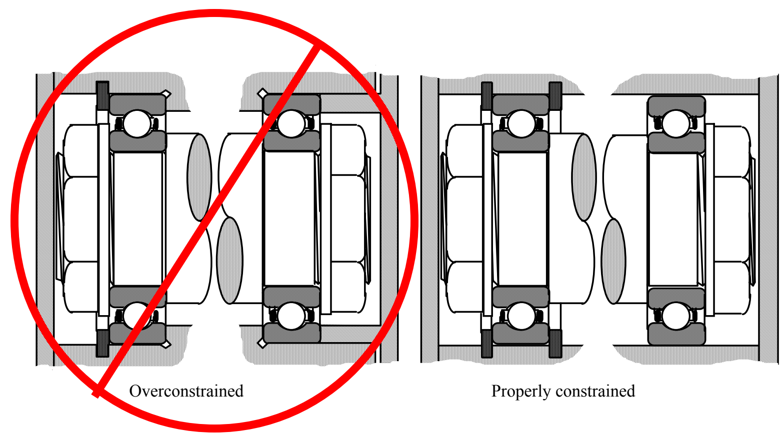
- Every rigid body has 6 *Degrees of Freedom* (DOF)
  - An exactly constrained design has no chance of deforming or having its function impaired be it by assembly, fastener tightening, thermal expansion, or external loads
- Make sure you have constrained what you want to constrain!
  - For a body to have N degrees of freedom free to move, there must be 6-N bearing reaction points!
  - To resist translation, a force is required.
  - To resist rotation, a moment, or two forces acting as a couple, is required!
    - Saint-Venant rules! Do not constrain a shaft with more than 2 bearings, unless it is very long…











Elastically Averaged Design

If *Exact Constraint Design*, has good and bad points, then Maxwell's reciprocity would indicate that *Inexact Constraint Design*, which would require systems to be statically indeterminate, might also have bad and good points! Indeed, a parameter that can make an over constrained 4 or 5 legged chair not wobble, is *controlled compliance*: Elastic deformation compensates for geometric errors; however, the forces are managed so yielding does not occur. When there are many compliant elements, each of which locally deforms to accommodate an error, in total they can form a very rigid and accurate system, and the design is called an *Elastically Averaged Design*.[1]

Carefully study the figure which shows the footprints of 3, 4, and 5 legged chairs. The red arrow shows the minimum radius from the center-of-stiffness, in this case the center of the chair, to the edge of the supports. This radius is indicative of the stability of the chair. If the center-of-mass shifts outside this radius' point, the chair will tip. Thus despite the fact that all three chairs have the same radius circle that contains all the legs, the more the legs, the greater the stability. However, in order to prevent the chair from rocking back and forth on three legs, because not all the legs' feet can ever lie in the exact same plane even if the floor was perfectly flat, the legs have to be compliant enough so that a modest load causes them to deflect and make them all contact the floor. On the other hand, the legs cannot be so soft that the sitter feels unstable when planting their mass onto the chair. Thus in order to design a five legged chair, the engineer has to have an idea of the potential variance in the floor and chair-leg planarity and the weight of the person?

An elastically averaged design can also support a system at many points, thus preventing large deformations sometimes associated with supporting a system at only a few points. Examples include large machine tools and manufacturing equipment which can have dozens of support points. But beware, these machines require very thick specially prepared concrete slabs, or else deformations in the foundations will cause machine deformations. This is also true of machine assemblies, where overall structural deformations can cause deformation of critical components unless they are properly supported.

1. John Strong, "New Johns Hopkins Ruling Engine", 1951 in the Journal of the Optical Society, Vol 41, pages 3ff

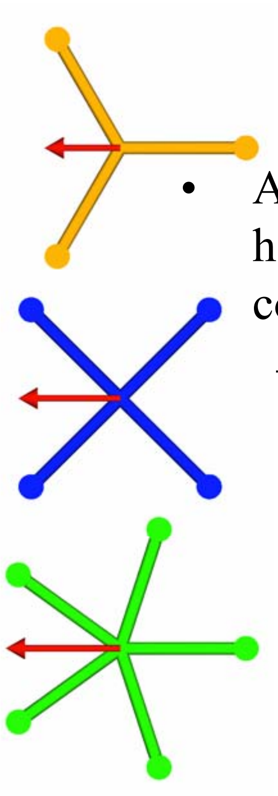
Consider a spoked bicycle wheel where any one spoke has marginal accuracy, strength, and stiffness, but together tensioned they create a stable and accurate wheel that is able to absorb many hazards and then keep rolling. Any spoke that becomes damaged will cause only a slight deformation in the wheel. Preloaded bearings and leadscrews also have all their individually compliant elements preloaded so they all make contact, and together yield high stiffness. Also consider the common gecko lizard which can stick to virtually any surface because the hairs on its feet subdivide into *seate*, which further subdivide into *spatulae* which make atomic contact with virtually any surface and enable Van der Waals forces to take hold![2]

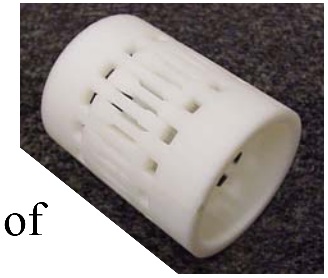
Many machine components are made far more accurate than any of their components by *elastic averaging*: *Hirth* or *Curvic* couplings use essentially two face gears that are forced together to allow one surface to be indexed with respect to another and achieve an accuracy (number of gear teeth)$^{1/2}$ better than either gear itself, because of the high forces used to preload the gears together. Spline-type flexible couplings can eliminate backlash if their elements radially flex to create a preload effect[3]. Hydrostatic or aerostatic bearings use pressurized fluid to support systems with 10-20x greater accuracy than any of the components with which they are made! In the limit, a relatively rigid body resting on a relatively soft body can be modelled as a beam on an elastic subgrade. Indeed, this is how concrete roadways are designed[4].

It is wise to not generalize: keep your eyes on the prize. What are the system functional requirements? How can you maximize total system performance and minimize cost? How can you create a design with the ability to evolve over time so you keep off the bleeding edge, stay on the leading edge, and stay ahead of the coagulated edge (see page 3-2 for a quick refresher)?

What aspects of your ***concepts*** are best made kinematic, and what aspects are best made elastically averaged? Carefully consider both options!

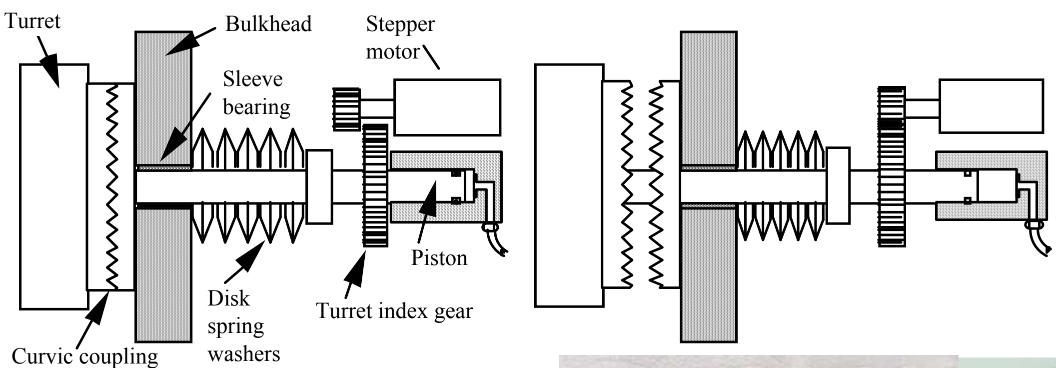
2. K. Autumn, Y. Liang, W.P. Chan, T. Hsieh, R. Fearing, T.W. Kenny, and R. Full, Dry Adhesive Force of a Single Gecko Foot-Hair, Nature. 405:681-685 (2000).
3. M. Balasubramaniam, H. Dunn, E. Golaski, S. Son, K. Sriram, A. Slocum, "An anti backlash two-part shaft with interlocking elastically averaged teeth", Precis. Eng., Volume 26, No. 3, pp. 314-330, 2002.
4. See R. J. Roark and W. C. Young, Formulas for Stress and Strain, 5th edition, McGraw-Hill Book Co., New York, 1975, p. 134; and S. Timoshenko Strength of Materials, Part II, 3rd ed., Robert E. Krieger Publishing Co., Melbourne, FL, p. 17.



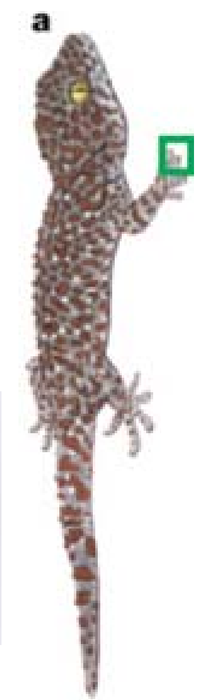


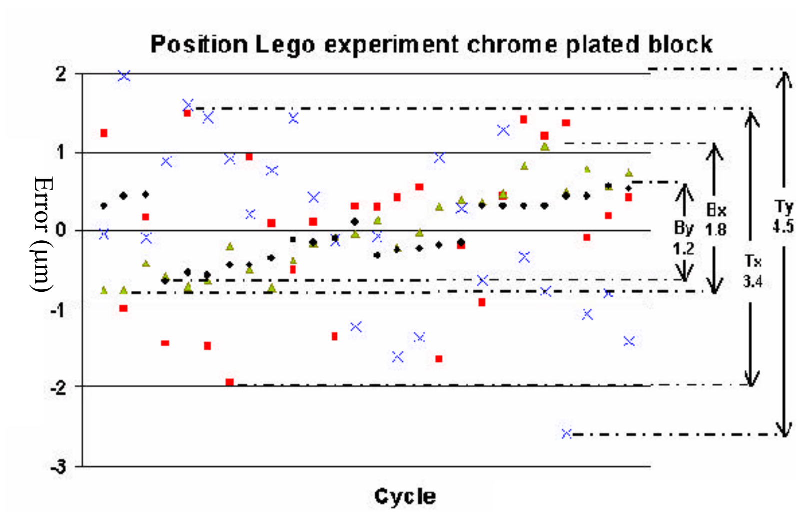
Elastically Averaged Design

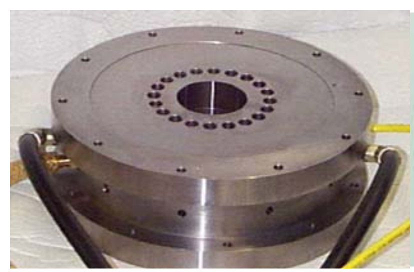
- Applying *Reciprocity* to *Exact Constraint Design* implies that instead of having an exact number of constraints, have an “infinite” number of constraints, so the error in any one will be averaged out!
  - Legos™, five legged chairs, windshield wipers, and Geckos are the most common examples, and many machine components achieve accuracy by elastic averaging

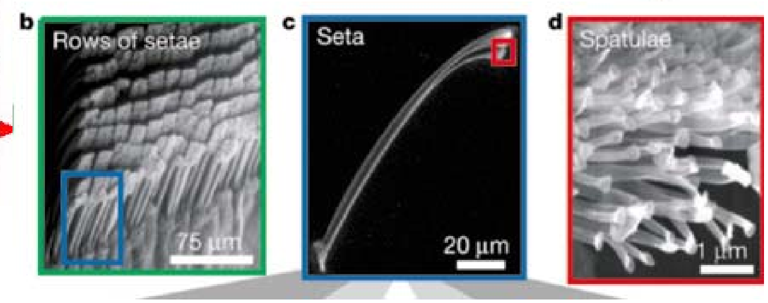


K. Autumn, Y. Liang, W.P. Chan, T. Hsieh, R. Fearing, T.W. Kenny, and R. Full, *Dry Adhesive Force of a Single Gecko Foot-Hair*, Nature. 405: 681-685 (2000)









Stick Figures

Design engineers sometimes trip over their own minds when running to create new ideas. They sometimes never finish the race because they become so tangled in their own thoughts. They sometimes run into walls because they cannot see through the clutter of all the possibilities. Hence they could often use a simple map to help guide them. They could often use a simple map that would allow them to see different possible routes, so they could pick the simplest one. They could use a method to record just the basic idea while capturing the fundamental principles, the keys to the idea, that make the idea one to change the world.

A *stick figure* can address these needs because it can be the simplest embodiment of one's thoughts. It can also serve as a tool that is used to set the stage for analysis of a design. It can have forces labeled and it can capture the designer's intent. Because of its inherent simplicity, it can be the quickest way for a mind to record what it is thinking so others can also understand and join in on the thinking fun! A stick figure fulfills this same function for the designer, letting the designer record just enough information to hold a thought, so the designer can reset the neurons and think new creative thoughts.

Stick figures also form the essence of *Free-Body-Diagrams*, which were discussed earlier in the context of applying Newton's Laws. Stick Figures enable the designer to sketch just what is needed to capture just enough geometry to enable them to visualize the forces and moments between components. Creating stick figures can be thought of as using Occam's razor to cut away all the non-essential details to reveal just what is needed!

How could a simple line drawing be a fundamental principle? How could a bunch of lines scribbled on a piece of paper rank way up there with Occam, Saint-Venant, Pythagoras (and Donald Duck!), Abbe, and Maxwell? The answer is very simple, the bunch of lines scrawled on a piece of paper by the hand of a designer are the very essences of Occam, Saint-Venant, Pythagoras, Abbe, and Maxwell. When a designer creates with their brain and passion fully engaged, they make all that the great minds sought to discover come to life. Their passion becomes forever immortalized, as it is passed forward to the next generation.

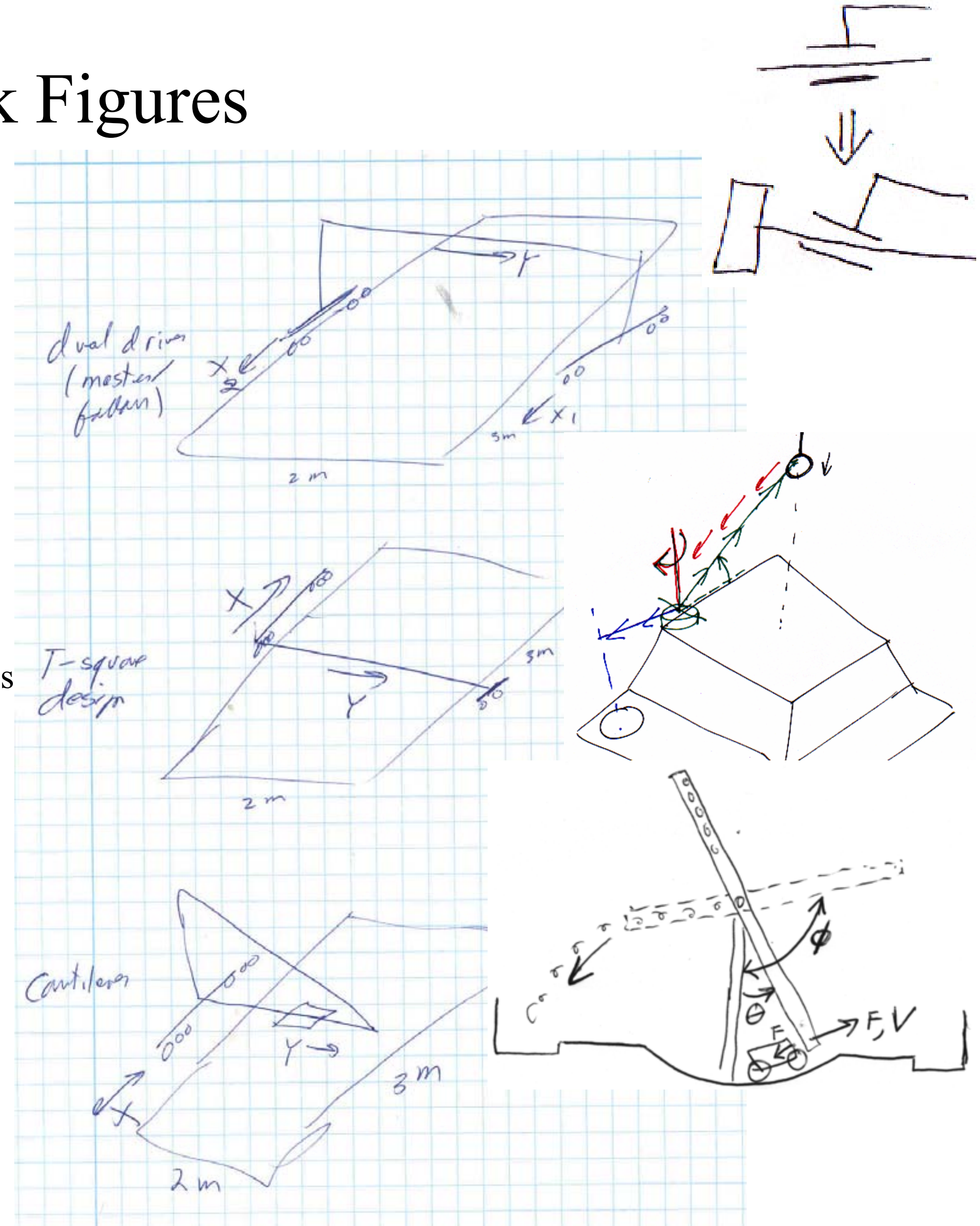
And what is the analytical equivalent of a stick figure? *Dimensional analysis!* Just the way a stick figure with lines allows a design engineer to frame a design figuratively, dimensional analysis allows an engineer to frame analysis of a problem. Remembering that calculations involving stress, for example, must result in units of N/m2, and energy equals the product of force and distance so it must have units of N-m, and power equals the product of force and velocity so it must have units of N-m/s, can save you immense grief!

With fundamental principles, great thoughts, passion, determination, and simple figures and appropriate analysis, ***strategies*** and ***concepts*** are born that can change the world.

Go forth and draw stick figures of your ***strategies*** and ***concepts*** while maintaining the proportions indicated by Saint-Venant, the Golden rectangle....and all the other FUNdaMENTAL principles!

Stick Figures

- Use of *fundamental principles* allows a designer to sketch a machine and error motions and coordinate systems just in terms of a *stick figure:*
  - The sticks join at centers of stiffness, mass, friction, and help to:
    - Define the sensitive directions in a machine
    - Locate coordinate systems
    - Set the stage for error budgeting
  - The designer is no longer encumbered by cross section size or bearing size
    - It helps to prevent the designer from locking in too early
- Error budget and preliminary load analysis can then indicate the required stiffness/load capacity required for each "stick" and "joint"
  - Appropriate cross sections and bearings can then be deterministically selected
- It is a "backwards tasking" solution method that is very very powerful!



Topic 3 Study Questions

Which suggested answers are correct (there may be more than one, or none)? Can you suggest additional and/or better answers?

1. An *error budget* for a machine should be created after all design layout is complete and all the major elements have been selected:
   - True
   - False
2. The purpose of an error budget is to:
   - Investigate the sensitivity of different design parameters (e.g., bearing accuracy classes) on system performance
   - Provide a quantitative comparison between design concepts
   - Tell you how much you can spend on reducing errors
   - Verify that chosen part tolerances are correct
3. A *sine error* δ, also known as the *Abbe effect,* is proportional to the product of the small angular error ε and the distance L from the source of the angular error to the point in question ($\delta = L\varepsilon$):
   - True
   - False
4. A "cosine error" δ, is proportional to the product of the small angular error ε squared and the distance L from the source of the angular error to the point in question ($\delta = L\varepsilon^2/2$):
   - True
   - False
5. A machine can be modeled with the use of a simple stick figure whose nodes are located at the centers of stiffness between components:
   - True
   - False
6. The more constraints placed on a component, the better its position will be defined:
   - True
   - False
7. The *center of stiffness* of a system is located at the most rigid location on the system:
   - True
   - False
8. The *center of friction* of a system is located at the point with the most friction in the system:
   - True
   - False
9. A *stick figure* is drawn using only simple lines to represent components and their connections, and it is used in the design of a machine to:
   - Show how the machine and machine operators might interact
   - Model the relative location of joints and axes to help with an early visualization of the machine and the preparation of an error budget for the machine.
   - Represent a machine without having to commit to specific components or structure
10. The terms *runout* and *error motion* are synonymous:
    - True
    - False
11. Every rigid body has six degrees of freedom:
    - True
    - False
12. *Exactly constraining* a component is *always* the best practice:
    - True
    - False
13. *Reciprocity* is used in structural analysis to indicate that if a force is applied at one point, and a deflection is observed at another point, then the force and deflections can be interchanged:
    - True
    - False
14. If your idea is not working, *reciprocity* can be used as a idea generator by:
    - Applying the loads at a different point
    - Conceptually turning your idea upside down or trying something completely different
    - Smashing your computer so it will never give you a bad answer ever again!
15. The *golden rectangle* is:

A rectangle whereby when a square is cut from it, the remaining rectangle has the same proportions (about 1.6:1)
Any rectangle whose proportions place the center of stiffness 2/3rds of the way towards one side of the rectangle
A tropical pastry

16. *Occam's razor* tells us that:
    If you are working really hard and not making real progress, there probably is a better way
    To cut the root of a problem, form a team and brainstorm before wasting too much time thinking by yourself
    Geeks should use electric shavers to keep from cutting themselves
17. *Saint-Venant's principle* says that:
    You should pray before the first time you hit the start button on the machine you just designed
    To control something, constrain it over several characteristic dimensions
    To not be constrained by something, be several characteristic dimensions away
18. *Characteristic dimensions* are:
    The most sensitive parameters describing the performance of a system
    The most critically toleranced dimensions on a drawing
    The smallest dimension of an object
    The largest dimension of an object
19. *Elastic averaging* is the principle of over constraining the system to maximize thermal stability:
    True
    False
20. *Elastic averaging* is used to help maximize the stiffness and repeatability of an interface:
    True
    False
21. *Elastic averaging* is achieved by:
    Providing for six constraints between two bodies, each of which has a stiffness greater than 1/6th the stiffness of the total system stiffness
    Providing for dozens of constraints, each of which has modest stiffness, but when preload is applied to the joint, they deform such that all are in contact thereby accommodating manufacturing tolerances.
    Putting on three pairs of socks
22. *Exact constraint design* is the principle of exactly constraining the system to only allow for the intended degrees of freedom:
    True
    False
23. *Exact constraint design* is used to help maximize the stiffness and repeatability of an interface:
    True
    False
24. *Exact constraint design* is achieved by:
    Providing for six constraints between two bodies, each of which typically has a stiffness on the order of 1/6th the stiffness of the total system stiffness
    Providing for dozens of constraints, each of which has modest stiffness, but when preload is applied to the joint, they deform such that all are in contact thereby accommodating manufacturing tolerances.
    Supporting a system with three elastic bands to minimize vibration pollution from external sources
25. The *structural loop* of a machine is the envelope that contains the entire structure of a machine:
    True
    False
26. The structural loop helps the designer determine:
    How the machine can be expected to deform as it moves from one position to another
    Where stiffness problems might likely occur
    Where thermal expansion problems might likely occur

Where symmetry might be applied to mitigate stiffness or thermal expansion problems

27. *Self help* is the principle of:
    - Using the internet to find answers to problems
    - Using a feature to balance, as opposed to resist by brute force, a load
    - Martha Stewart meeting Oprah Winfrey
28. Preload is used to remove gaps between components and prevent the occurrence of gaps by creating some nominally constant displacement of the inherent compliance in a system, particularly the bearings:
    - True
    - False
29. Preload can be very dangerous to a system because thermal expansion or geometric error in the components can cause overloading:
    - True
    - False
30. To prevent preload-induced overloading, the preload means must be able to accommodate thermal expansion or geometric errors in components without overloading the system, and this is typically achieved by making the preload mechanism compliance greater than the primary load path compliance:
    - True
    - False
31. Given the stiffness $K_{primary}$ of the primary load path and the stiffness $K_{preload}$ of the preload means, the total system stiffness is:
    - $K_{primary} + K_{preload}$
    - $1/(1/K_{primary} + 1/K_{preload})$
32. The principle of self-help can be used to balance pressure forces to minimize pressure-induced gap expansion:
    - True
    - False
33. Saint Venant's principle can help to determine optimal bolt spacing:
    - True
    - False
34. Geometric errors are the easiest to deal with because they typically can be mapped and compensated for in the control system:
    - True
    - False
35. Dynamic errors can be difficult to deal with because the system typically does not have the bandwidth with which to correct for them:
    - True
    - False
36. The first buckling mode load capacity of a simply supported shaft is approximately 2.47*E*I/Length**2:
    - True
    - False
37. The first natural frequency of a simply supported shaft is 1.875^2*(E*I/(Area*density*Length**4))
    - True
    - False
38. *Repeatability* is more important than accuracy because repeatability can be mapped and used to obtain accuracy by servo control:
    - True
    - False
39. Accuracy is the ability of a machine to "tell the truth":
    - True
    - False
40. *Repeatability* is the same as precision:
    - True
    - False
41. *Repeatability* is the ability of a machine to "tell the same story every time"
    - True
    - False
42. *Resolution* is the limit of a machine's ability to move a minimum distance:
    - True
    - False
43. A *metrology frame* is not subject to deformations caused by motions and loads placed upon the machine that it is used to measure:

True
False

44. The *principle of self-help* can be applied to use the attractive force between open-face iron-core linear electric motor components to preload aerostatic bearings:
    True
    False