Earth Science Applications of Space Based Geodesy

DES-7355

Tu-Th 9:40-11:05

Seminar Room in 3892 Central Ave. (Long building)

Bob Smalley
Office: 3892 Central Ave, Room 103
678-4929

Office Hours – Wed 14:00-16:00 or if I'm in my office.

http://www.ceri.memphis.edu/people/smalley/ESCI7355/ESCI_7355_Applications_of_Space_Based_Geodesy.html

Class 7

So far

Have not specified type of arrival.

Can do with P only, S only (?!), P and S together, or S-P.

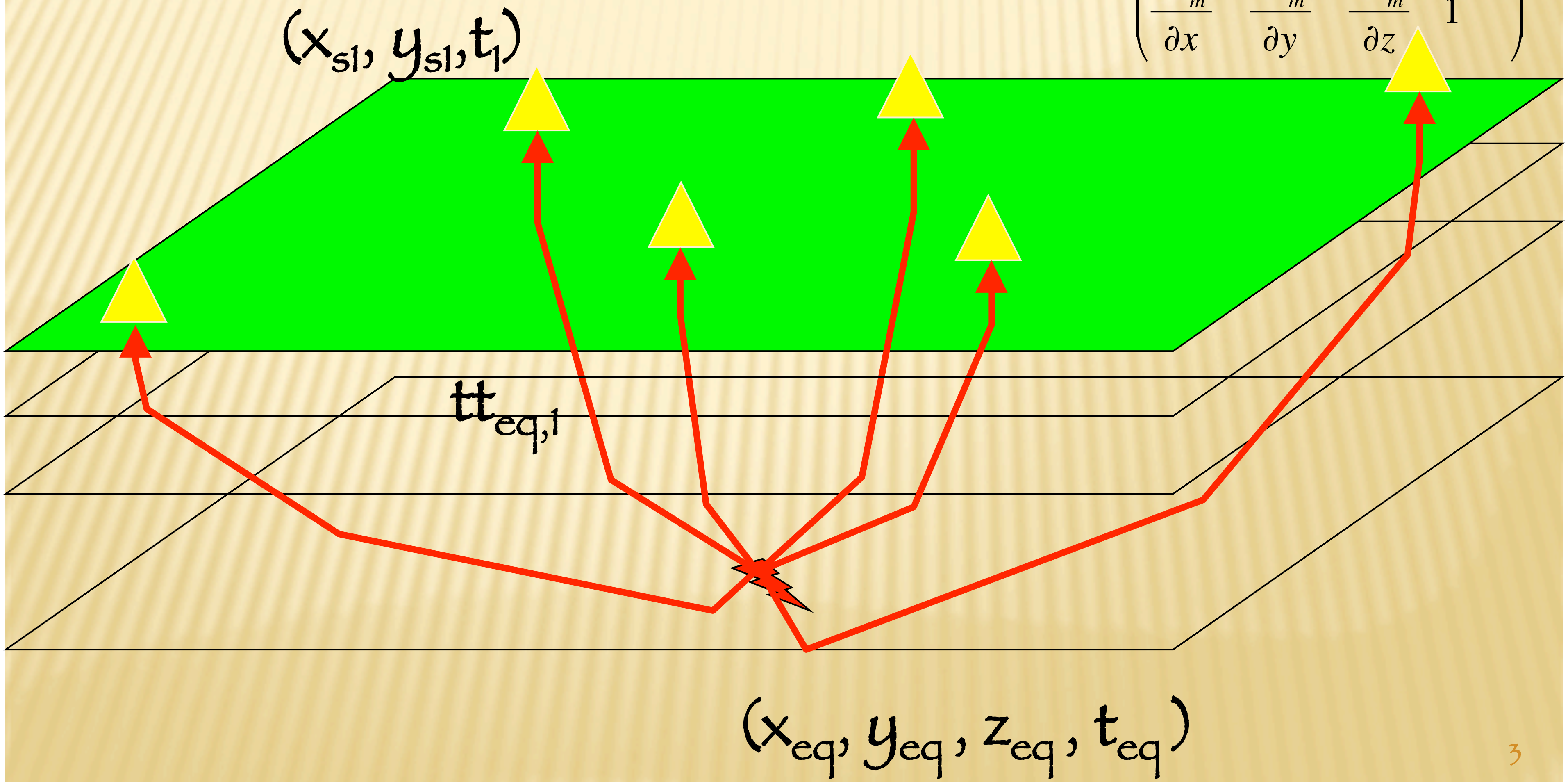
Need velocity <u>model</u> to calculate travel times and travel time derivatives
(so earthquakes are located with respect to the assumed velocity model, not real earth.
Errors are "formal", i.e. with respect to model.)

Velocity models usually laterally homogeneous.

Problems:

Column of 1's – if one of the other columns is constant (or approximately constant) matrix is singular and can't be inverted.

$$A = \begin{pmatrix} \frac{\partial \tau_1}{\partial x} & \frac{\partial \tau_1}{\partial y} & \frac{\partial \tau_1}{\partial z} & 1 \\ \frac{\partial \tau_2}{\partial x} & \frac{\partial \tau_2}{\partial y} & \frac{\partial \tau_2}{\partial z} & 1 \\ \frac{\partial \tau_3}{\partial x} & \frac{\partial \tau_3}{\partial y} & \frac{\partial \tau_3}{\partial z} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \tau_m}{\partial x} & \frac{\partial \tau_m}{\partial y} & \frac{\partial \tau_m}{\partial z} & 1 \end{pmatrix}$$

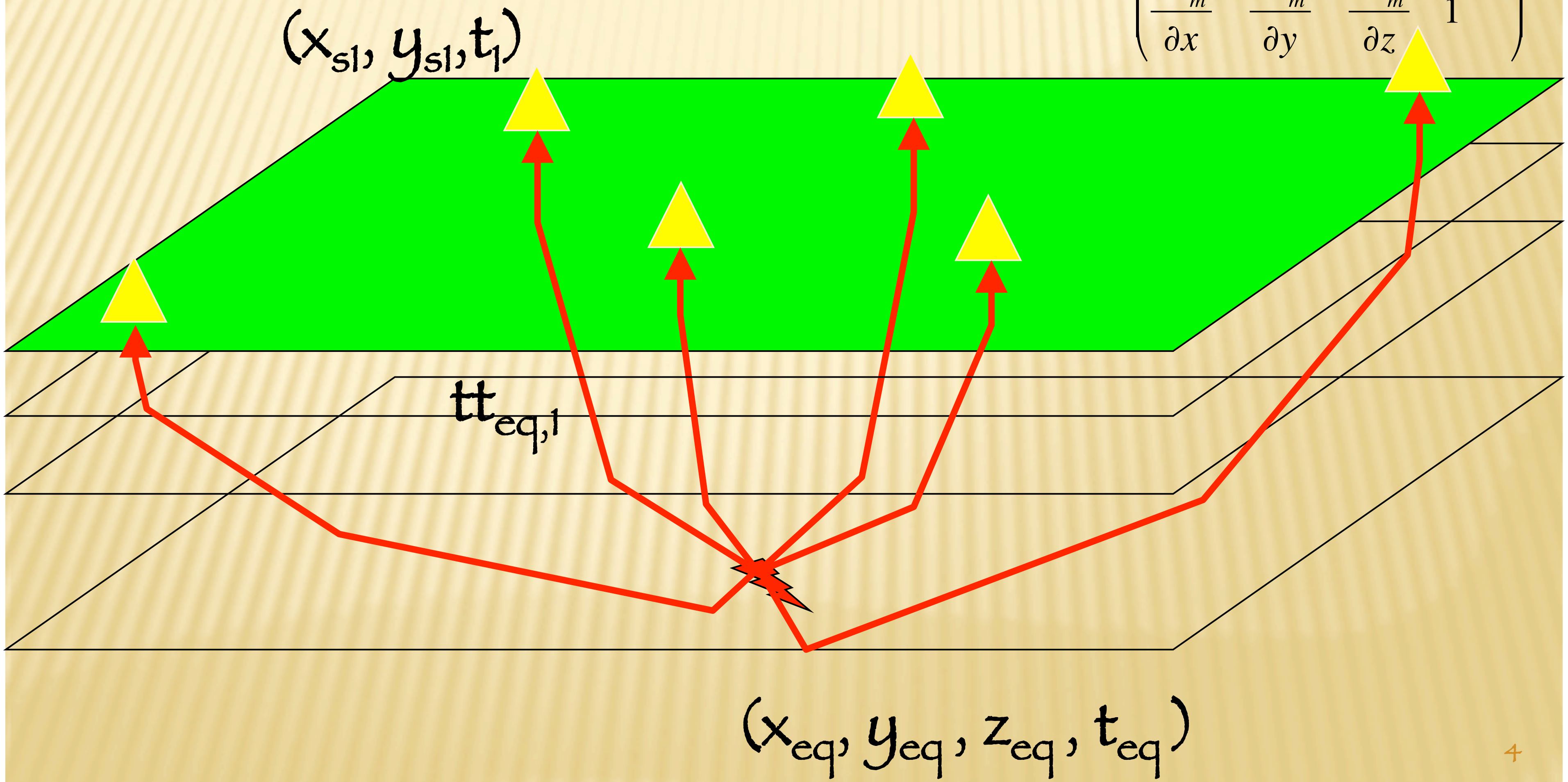


How can this happen:

-All first arrivals are head waves from same refractor

- Earthquake outside the network

$$A = \begin{pmatrix} \frac{\partial \tau_1}{\partial x} & \frac{\partial \tau_1}{\partial y} & \frac{\partial \tau_1}{\partial z} & 1 \\ \frac{\partial \tau_2}{\partial x} & \frac{\partial \tau_2}{\partial y} & \frac{\partial \tau_2}{\partial z} & 1 \\ \frac{\partial \tau_3}{\partial x} & \frac{\partial \tau_3}{\partial y} & \frac{\partial \tau_3}{\partial z} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \tau_m}{\partial x} & \frac{\partial \tau_m}{\partial y} & \frac{\partial \tau_m}{\partial z} & 1 \end{pmatrix}$$

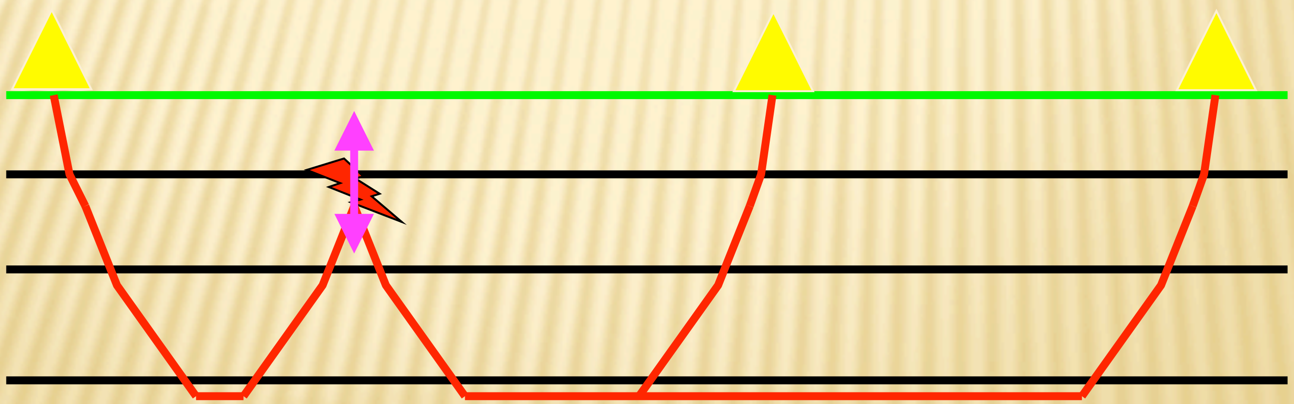


All first arrivals are head waves from same refractor

$$\frac{\partial \tau_k}{\partial z} = \text{constant } \forall\, k$$

In this case we cannot find the depth and origin time independently.

$$A = \begin{pmatrix} \frac{\partial \tau_1}{\partial x} & \frac{\partial \tau_1}{\partial y} & \frac{\partial \tau_1}{\partial z} & 1 \\ \frac{\partial \tau_2}{\partial x} & \frac{\partial \tau_2}{\partial y} & \frac{\partial \tau_2}{\partial z} & 1 \\ \frac{\partial \tau_3}{\partial x} & \frac{\partial \tau_3}{\partial y} & \frac{\partial \tau_3}{\partial z} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \tau_m}{\partial x} & \frac{\partial \tau_m}{\partial y} & \frac{\partial \tau_m}{\partial z} & 1 \end{pmatrix} = \begin{pmatrix} \frac{\partial \tau_1}{\partial x} & \frac{\partial \tau_1}{\partial y} & c & 1 \\ \frac{\partial \tau_2}{\partial x} & \frac{\partial \tau_2}{\partial y} & c & 1 \\ \frac{\partial \tau_3}{\partial x} & \frac{\partial \tau_3}{\partial y} & c & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \tau_m}{\partial x} & \frac{\partial \tau_m}{\partial y} & c & 1 \end{pmatrix}$$



Earthquake outside the network

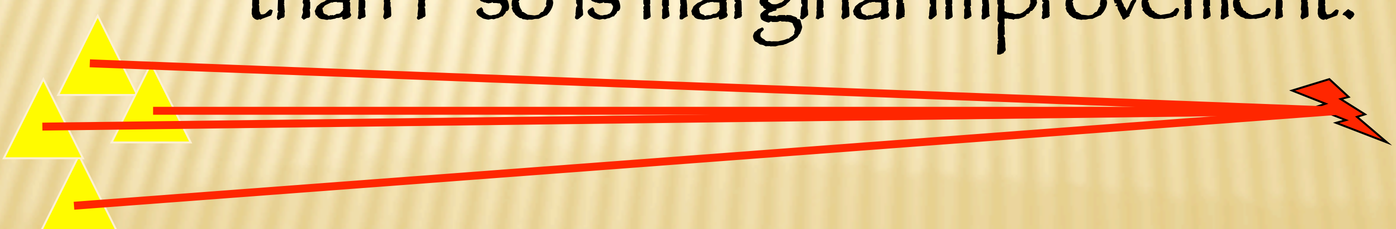
$$\frac{\partial \tau_k}{\partial x} \approx \text{constant } \forall\, k$$

$$\frac{\partial \tau_k}{\partial y} \approx \text{constant } \forall\, k$$

$$A = \begin{pmatrix} \frac{\partial \tau_1}{\partial x} & \frac{\partial \tau_1}{\partial y} & \frac{\partial \tau_1}{\partial z} & 1 \\ \frac{\partial \tau_2}{\partial x} & \frac{\partial \tau_2}{\partial y} & \frac{\partial \tau_2}{\partial z} & 1 \\ \frac{\partial \tau_3}{\partial x} & \frac{\partial \tau_3}{\partial y} & \frac{\partial \tau_3}{\partial z} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \tau_m}{\partial x} & \frac{\partial \tau_m}{\partial y} & \frac{\partial \tau_m}{\partial z} & 1 \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & \frac{\partial \tau_1}{\partial z} & 1 \\ c_1 & c_2 & \frac{\partial \tau_2}{\partial z} & 1 \\ c_1 & c_2 & \frac{\partial \tau_3}{\partial z} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ c_1 & c_2 & \frac{\partial \tau_m}{\partial z} & 1 \end{pmatrix}$$

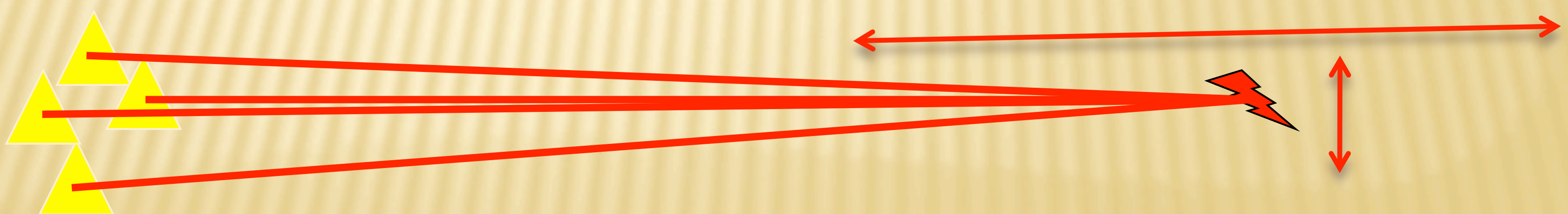
In this case only the azimuth is constrained.

If using both P and S, can also get range, but S "noisier" than P so is marginal improvement.



Probably also suffering from depth-origin time coupling

Problem gets worse with addition of noise (changes length of red lines – intersection point moves left/ right – change of distance - much more than in perpendicular direction – change of azimuth.)

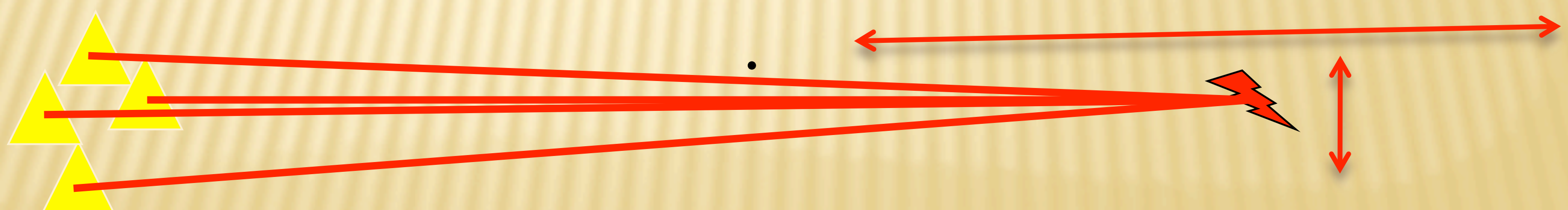


Similar problems with depth.

d/dz column ~equal, so almost linearly dependent on last column

and

gets worse with addition of noise (changes length of red lines – intersection point moves left/right [depth, up/down {drawn sideways}] much more than in perpendicular direction [position].)



Other problems:

Earthquake locations tend to “stick-on” layers in velocity model.

When earthquake crosses a layer boundary, or the depth change causes the first arrival to change from direct to head wave (or vice verse or between different head waves), there is a discontinuity in the travel time derivative (Newton’s method). May move trial location a large distance.

Solution is to “damp” (limit) the size of the adjustments – especially in depth.

Other problems:

Related to earthquake location, but bigger problem for focal mechanism determination.

Raypath for first arrival from solution may not be actual raypath, especially when first arrival is head wave.
Results in wrong take-off angle.

Since head wave usually very weak, oftentimes don't actually see head wave. Measure P arrival time, but location program models it as P_n.

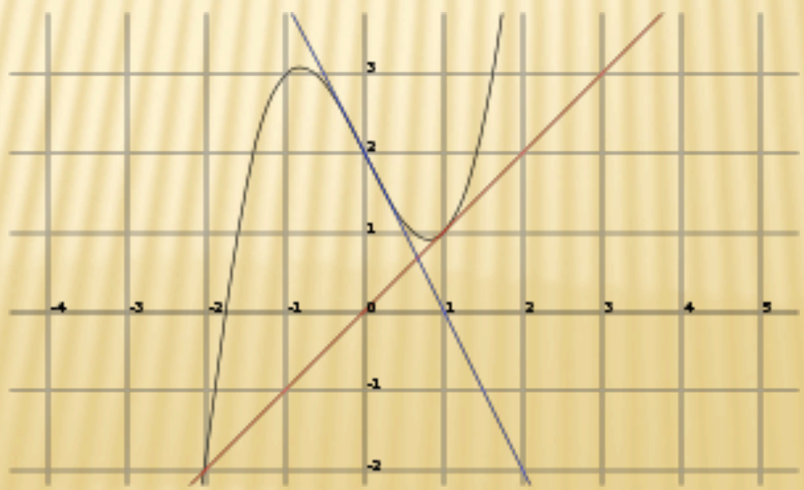
A look at Newton's method

Want to solve for zero(s) of $F(x)$

Start with guess, x_0.

Calculate $F(x_0)$ (probably not zero, unless VERY lucky).

Find intercept $x_1 = x_0 - F(x_0)/F'(x_0)$



Newton's method

Want to solve for zero(s) of $F(x)$

Now calculate $F(x_1)$.

See how close to zero.

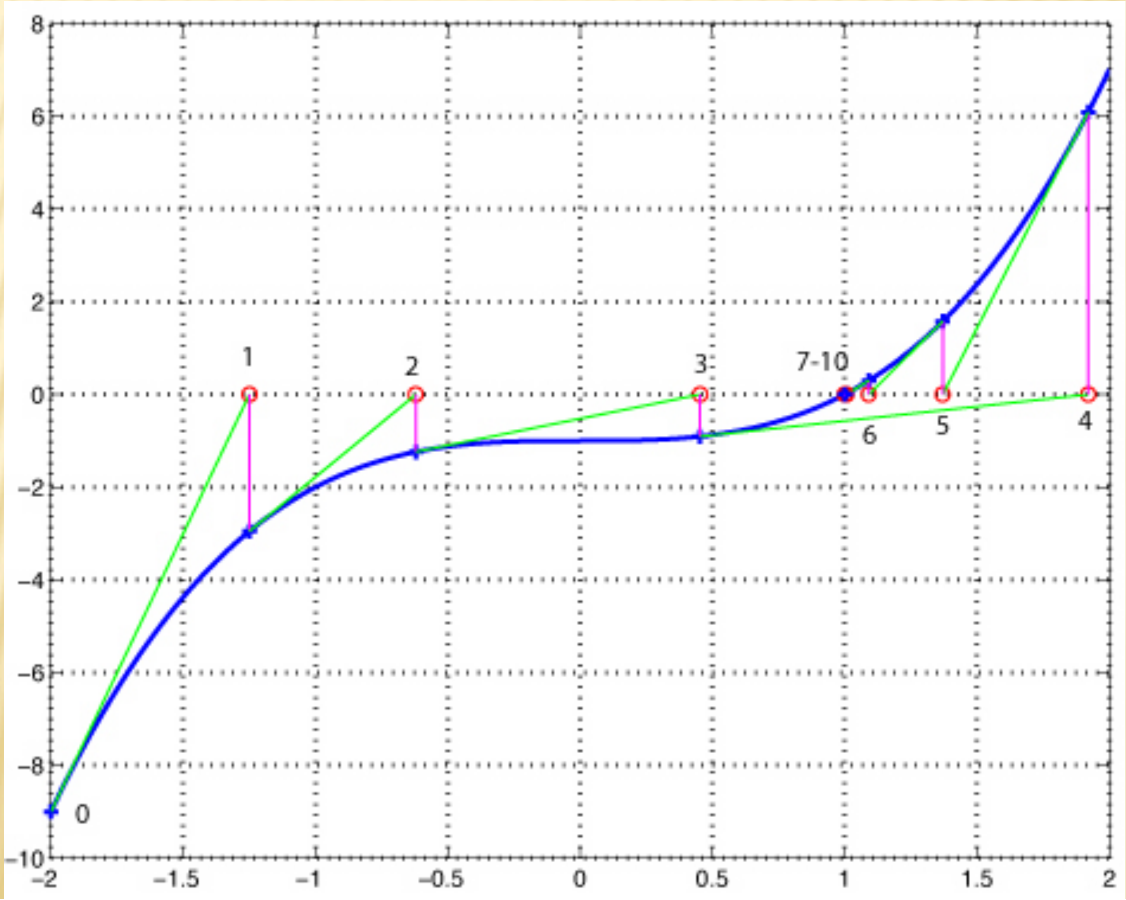
If close enough – done.

Newton's method

If not "close enough", do again

Find intercept $x_2 = x_1 - F(x_1)/F'(x_1)$

If close enough, done, else – do again.

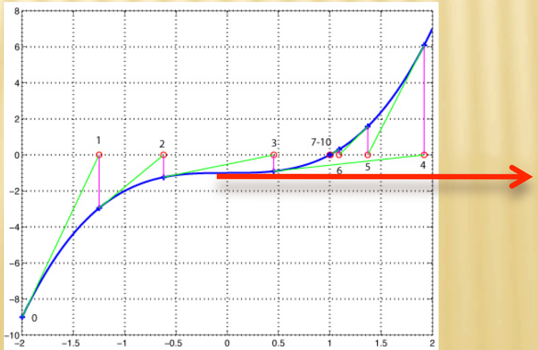


Newton's method

$$X_{n+1} = x_n - F(x_n)/F'(x_n)$$

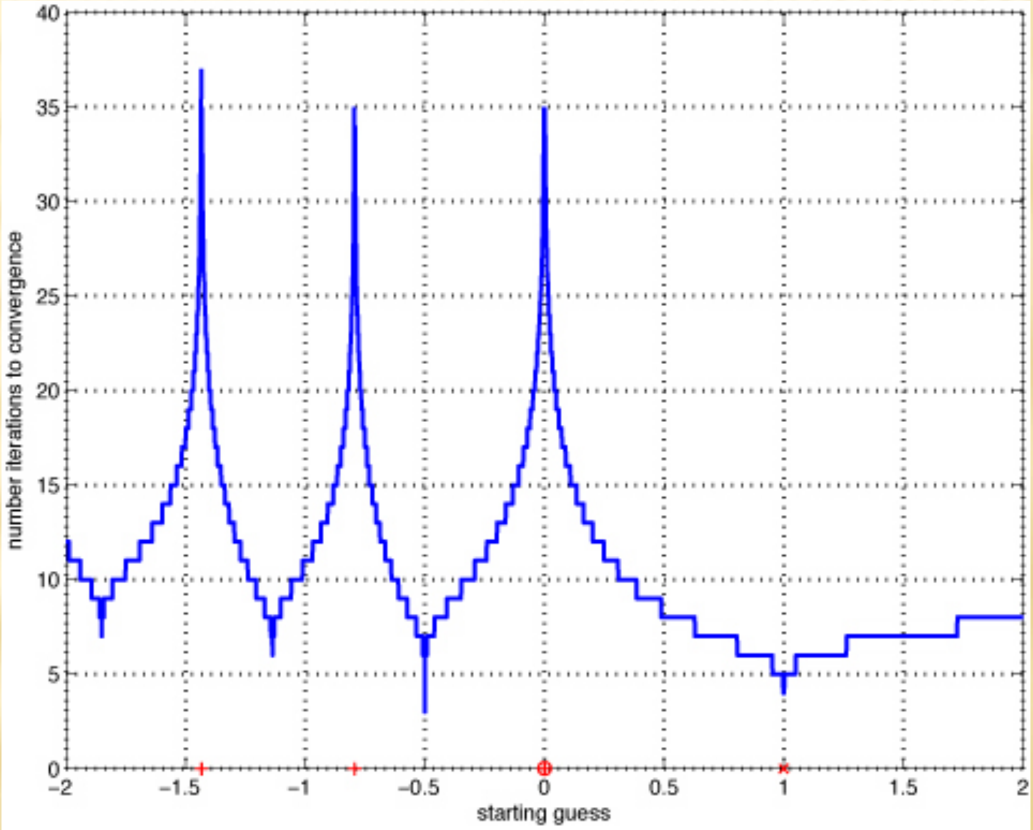
What happens when $F'(x_n)=0$?

Geometrically, you get sent off to infinity – method fails.
(Mathematically can't divide by zero – method fails.)



Newton's method

How does convergence depend on starting value?



Some starting values iterate through x_n=0 and therefore do no converge (limited calculation to 35 iterations).

Newton's method

Other problems

Point is "stationary" (gives back itself x_n -> x_n…).

Iteration enters loop/cycle: $x_n \to x_{n+1} \to x_{n+2} = x_n \ldots$

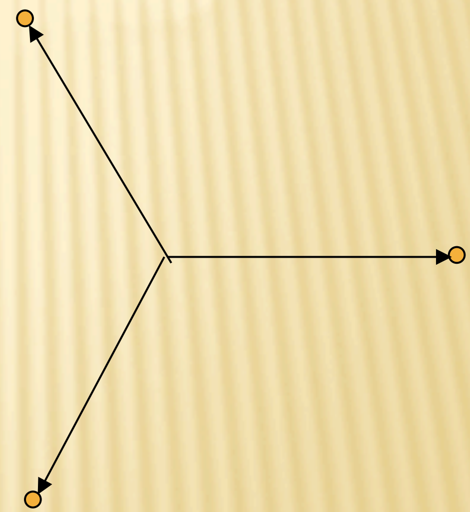
Derivative problems (does not exist).

Discontinuous derivative.

Newton's method applied to solution of non-linear, complex valued, equations

Consider

$Z^3 - 1 = 0.$



Newton's method applied to solution of non-linear, complex valued, equations

Consider

$Z^3-1=0.$

Solutions

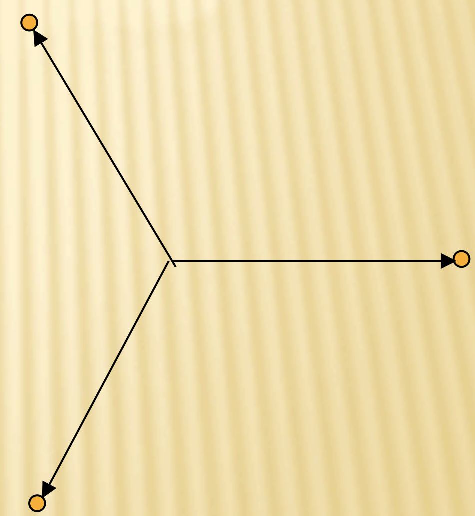
Three of them

$1\, e^{(i2\pi n/3)}$

n=0, 1, 2

Distance = 1

Every 120 degrees



Newton's method applied to solution of non-linear, complex valued, equations

Consider

$Z^3-1=0$

Solutions

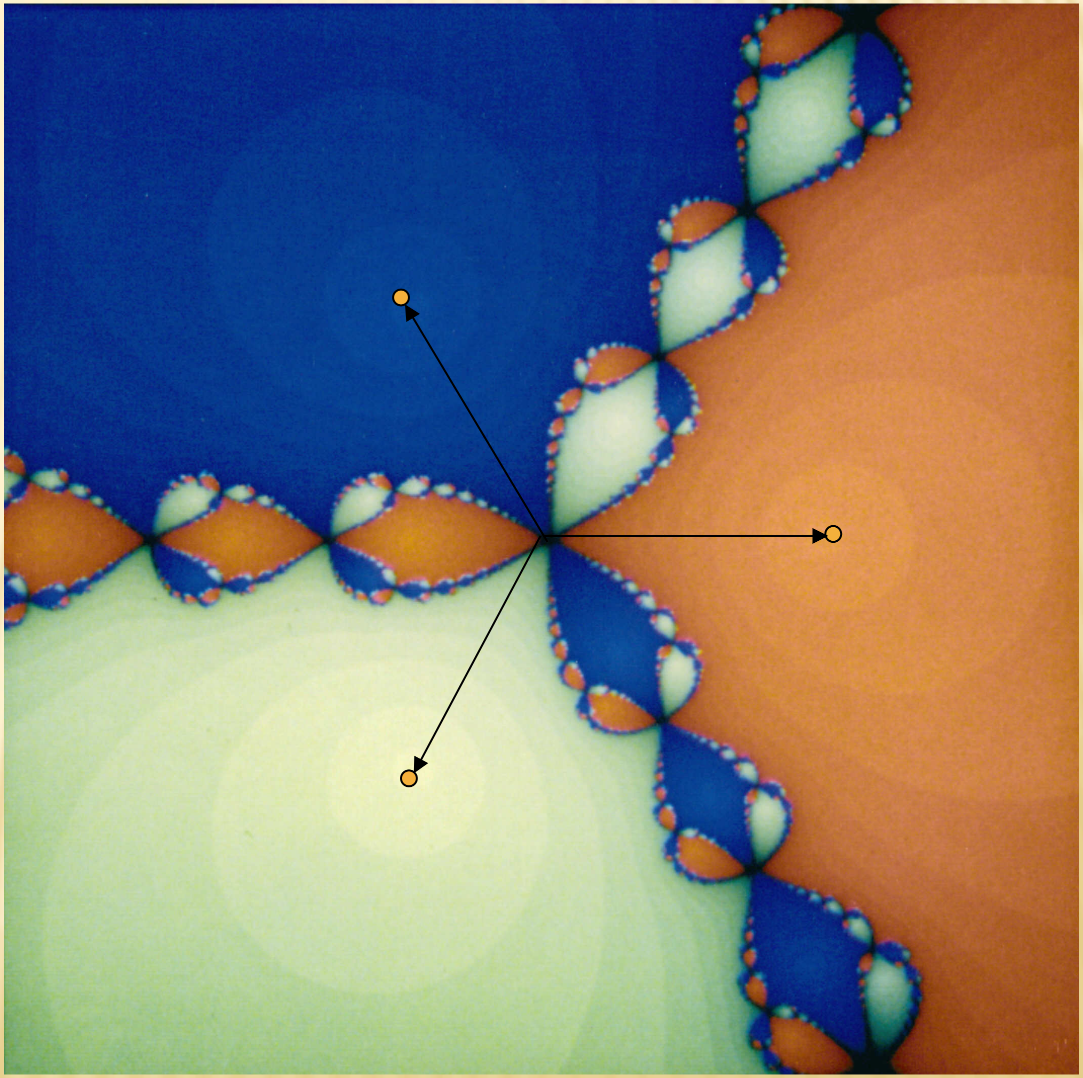
Three of them

$1\, e^{(i2\pi n/3)}$

n=0, 1, 2

Distance = 1

Every 120 degrees

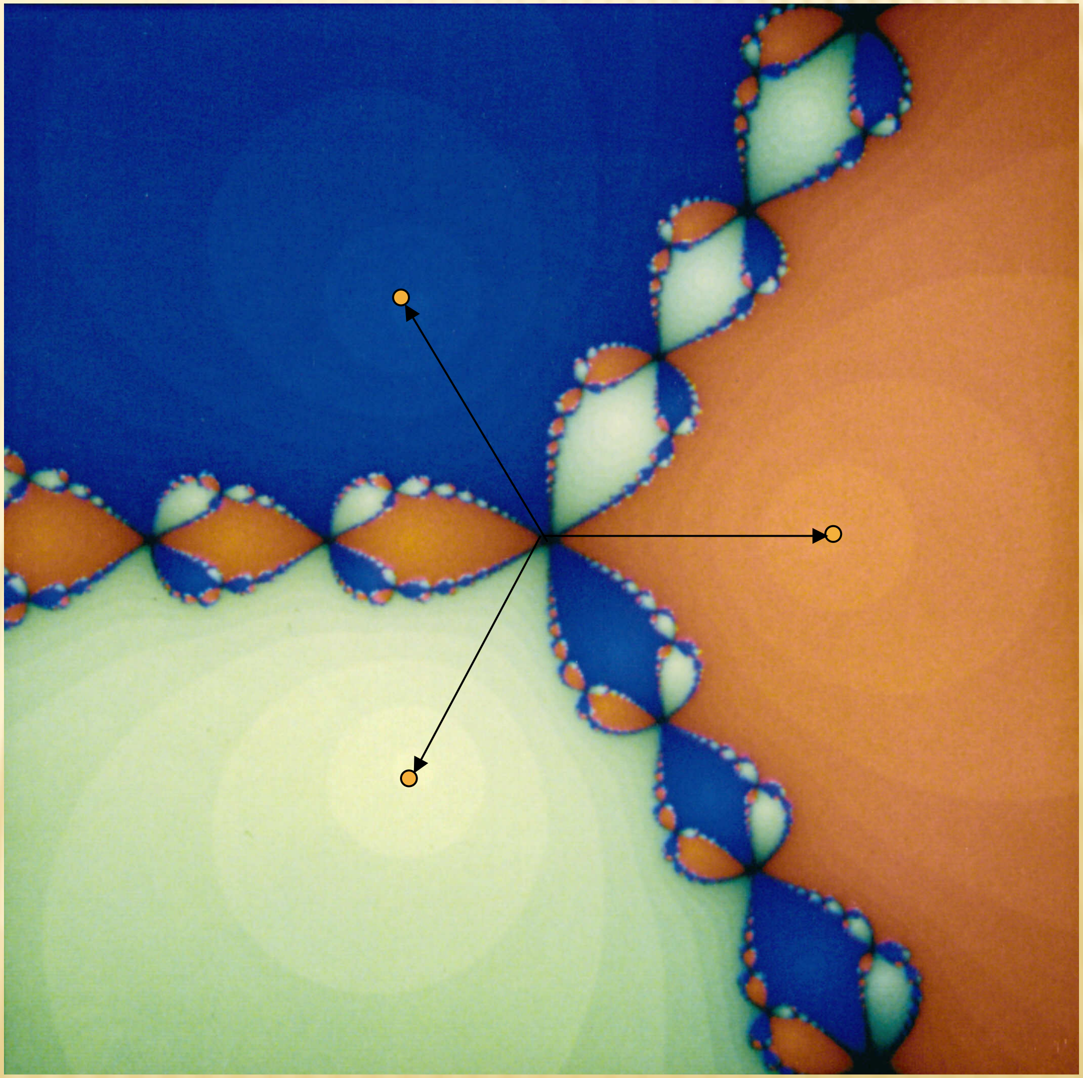


Take each point in the complex plane as a starting guess and apply Newton's method.

Now

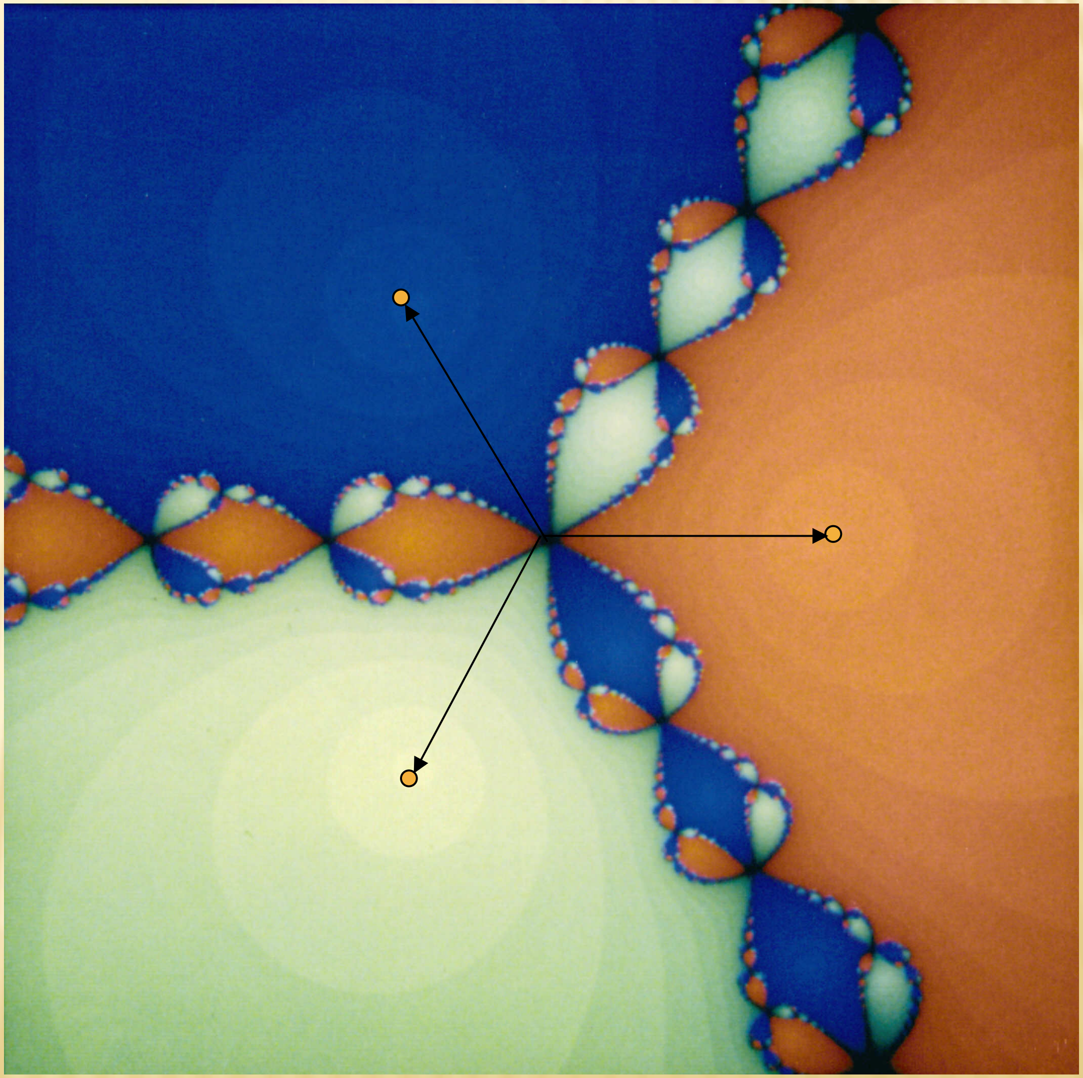
Color the starting points to identify which of the three roots each starting point converges to using Newton's method.

eg. all the red points converge to the root at 1.



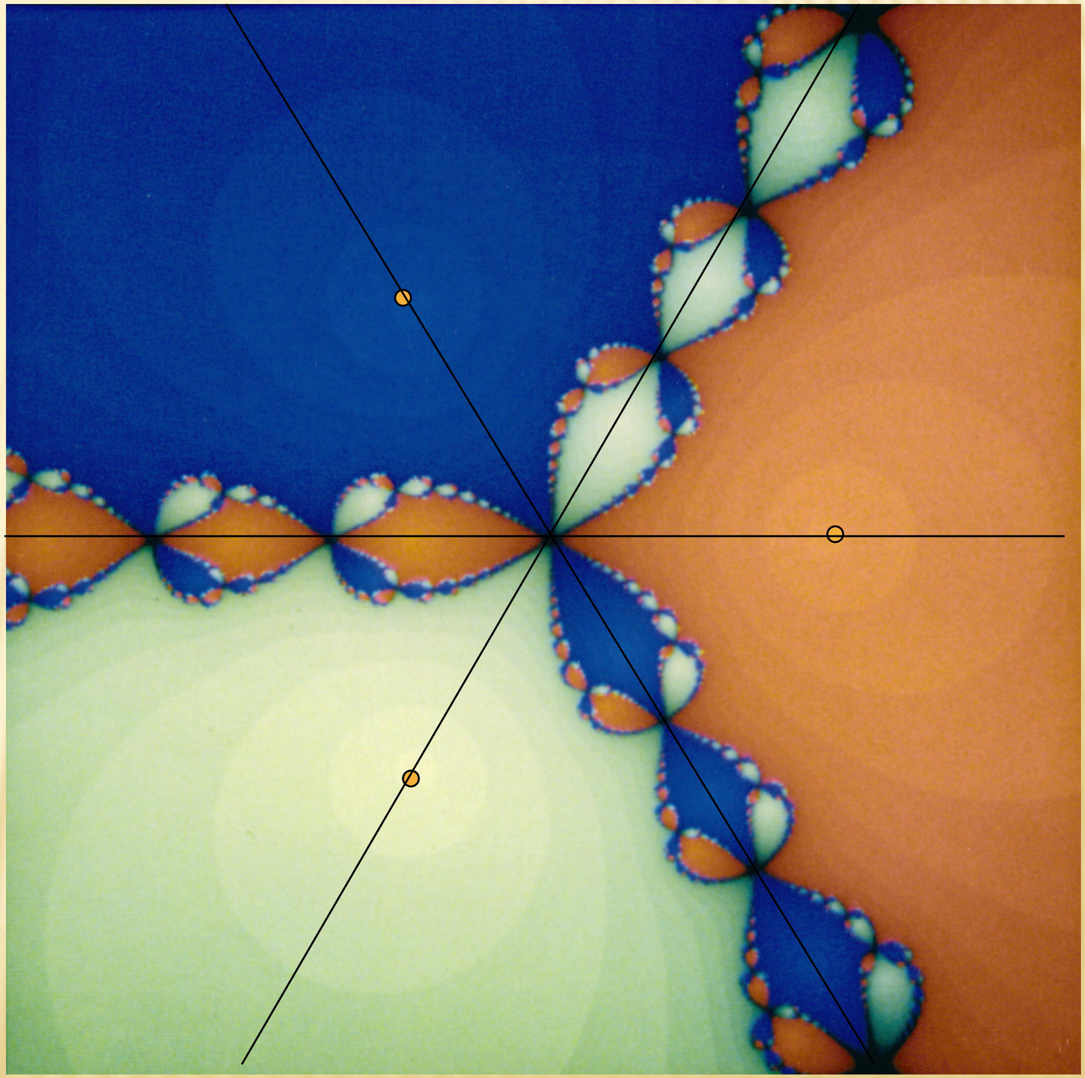
Let the intensity of each starting point's color be related to the number of steps to converge to that root

(brighter - converges faster, darker – converges slower)

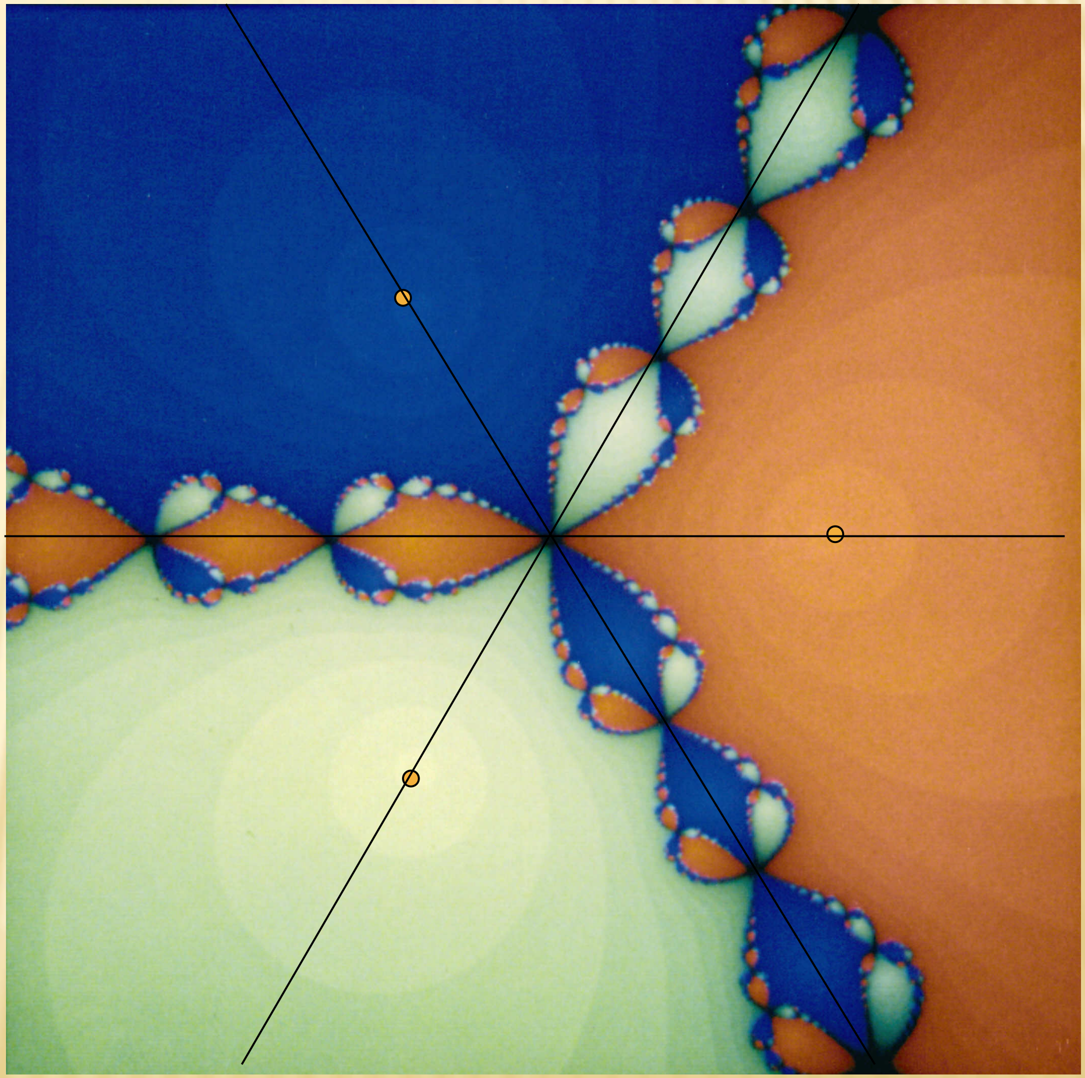


Notice that any starting point on the real line converges to the root at 1

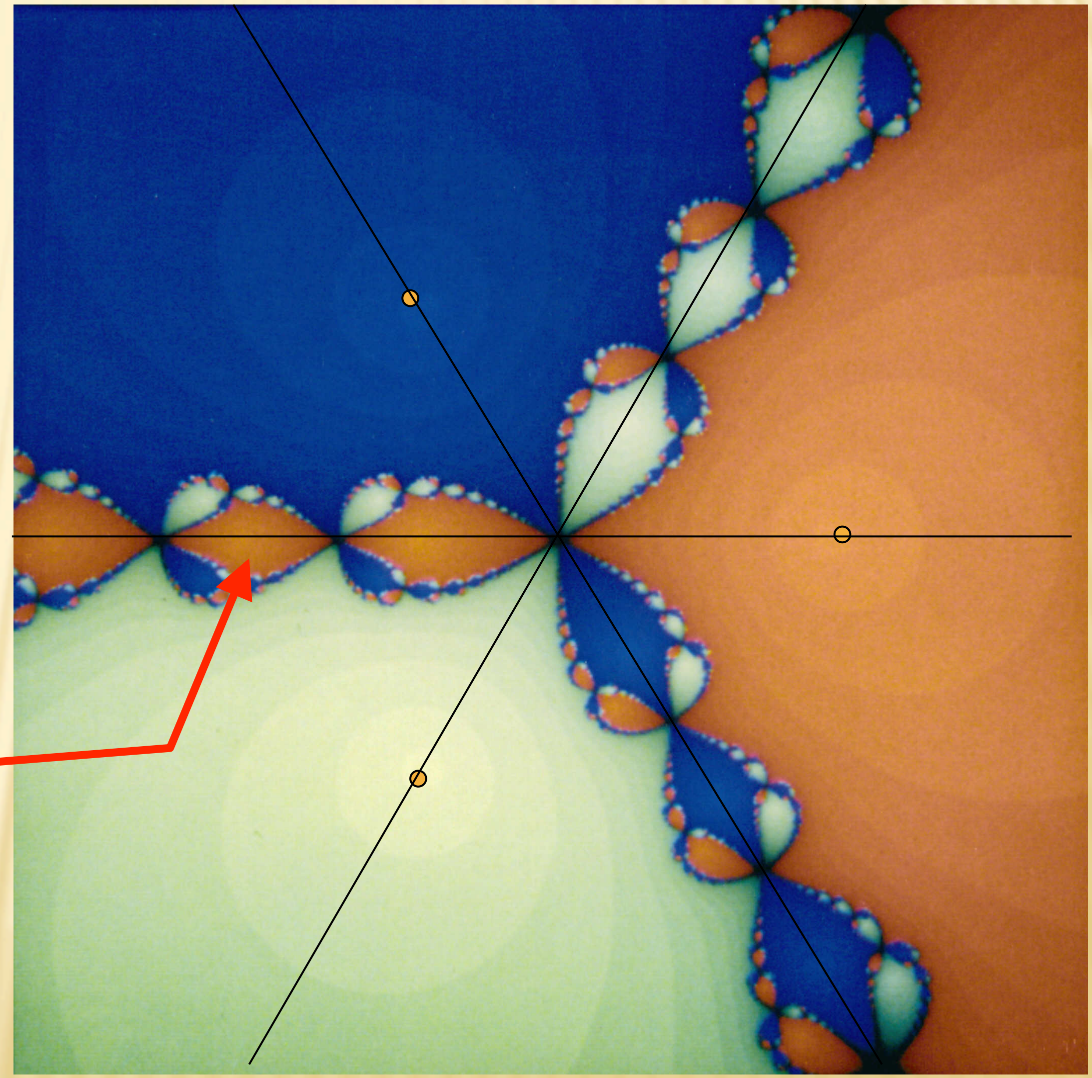
Similarly points on line sloping 60 degrees converge to the other 2 roots.



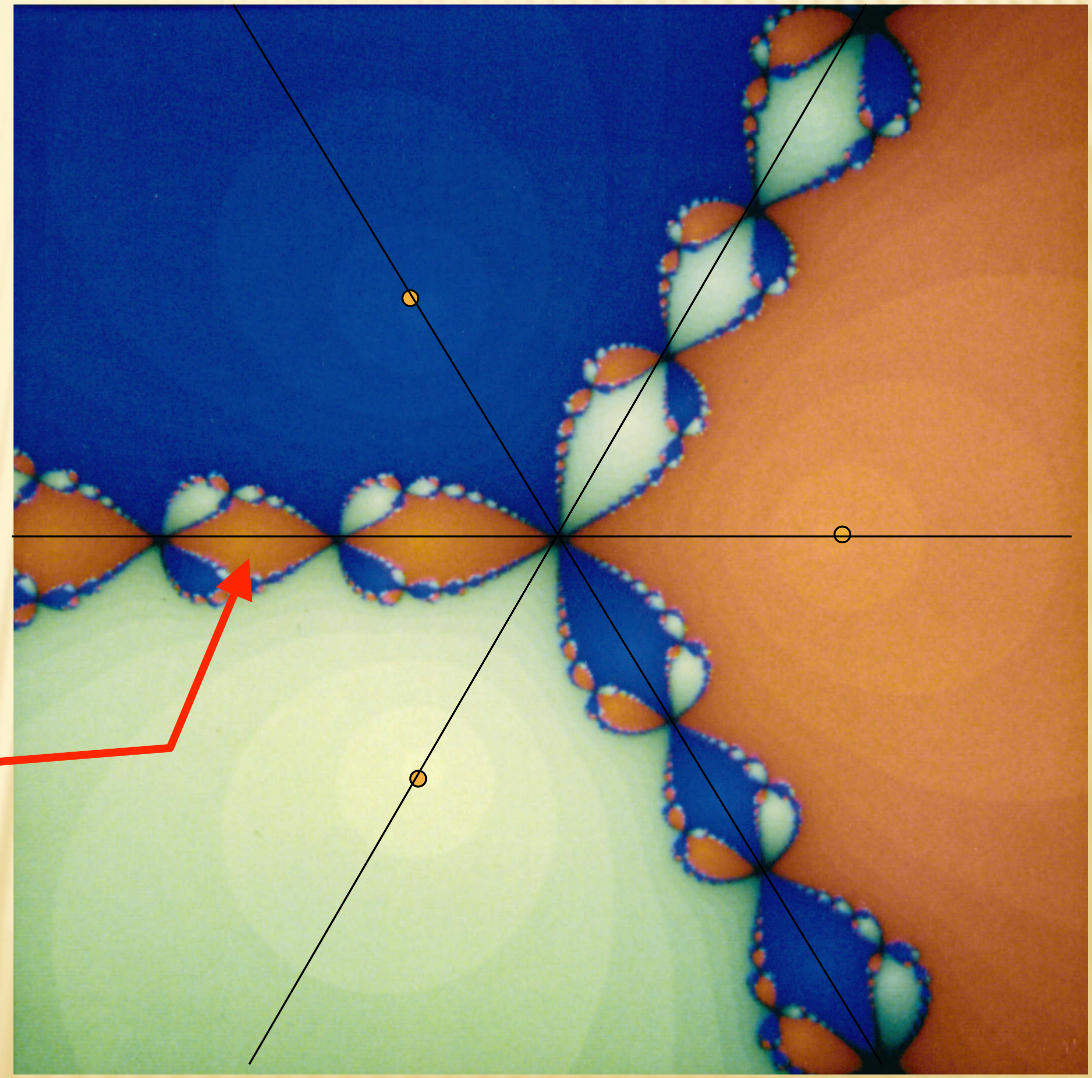
Notice that in the ~third of the plane that contains each root things are pretty well behaved.



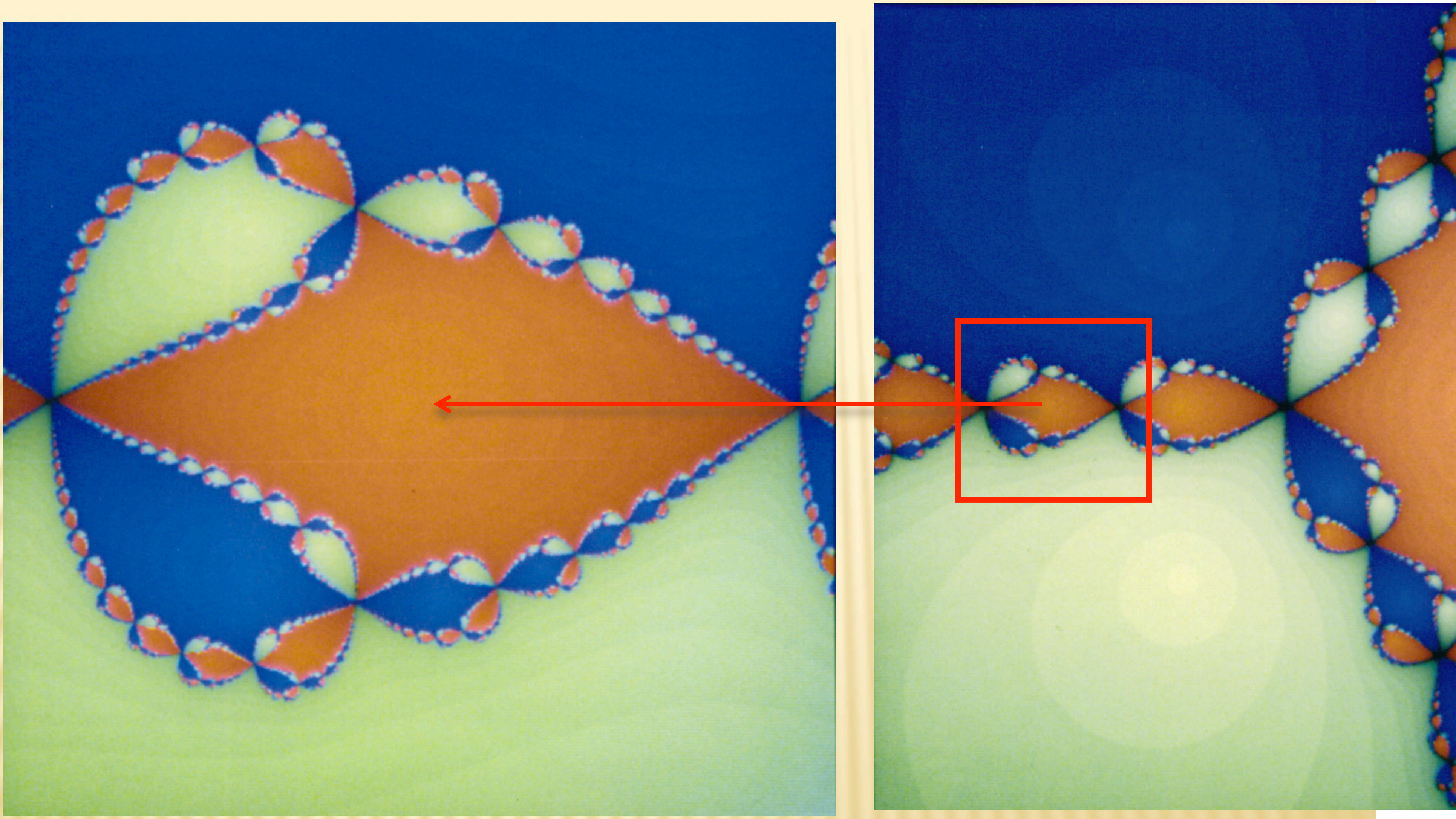
Notice that where any two domains of convergence meet, it looks a little complicated.



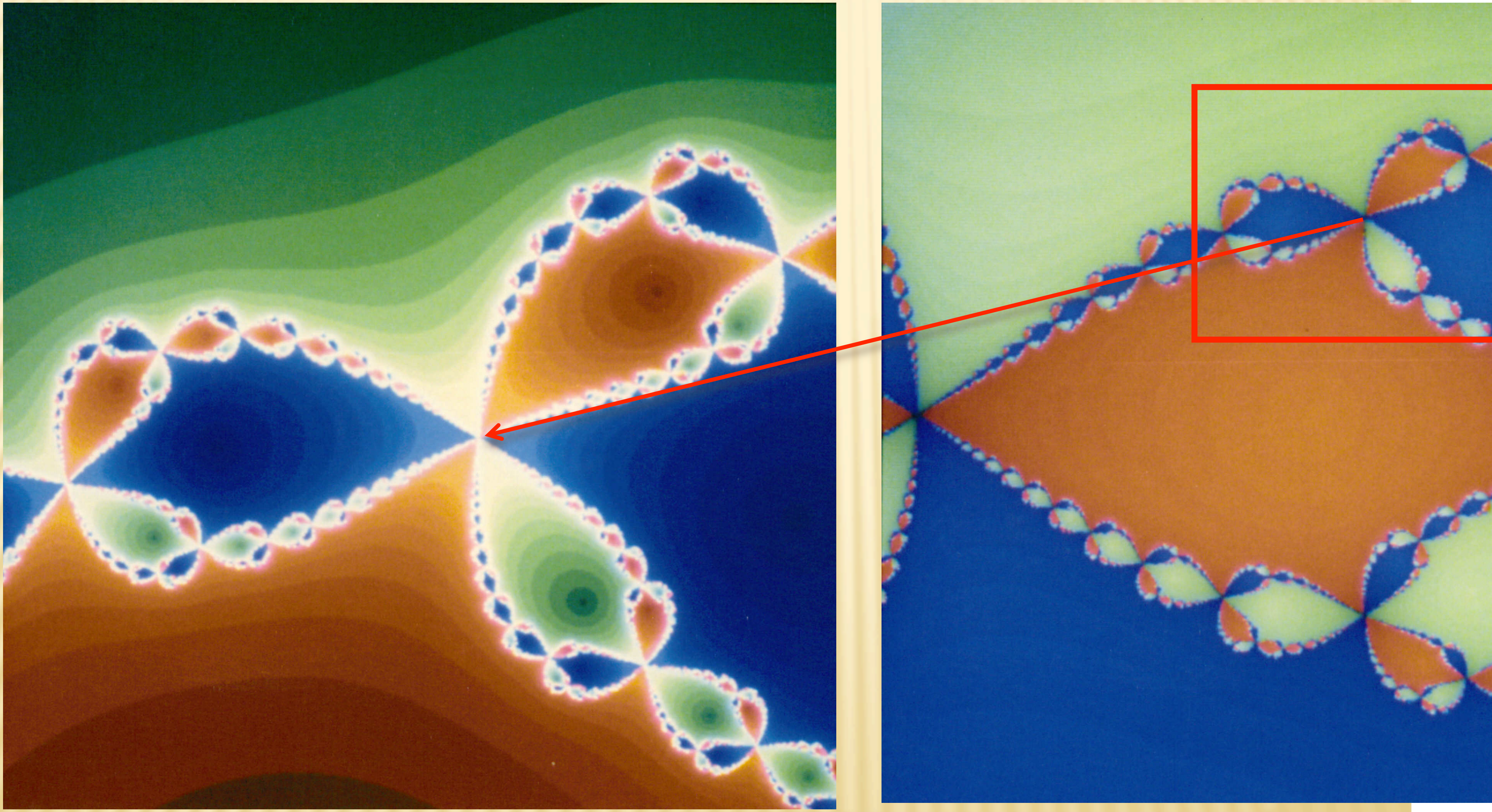
Basically – the division between any two colors is always separated by the third color.



AT ALL SCALES!



Zoom in



Zoom in again

If you keep doing this (zoom in) the "triple" junctions start to look like

Mandlebrot sets!

and you will find points that either never converge or converge very slowly.

Quick implication –

linear iteration to solve non-linear inversion problems

(Newton's method, non-linear least squares, etc.)

may be unstable.

More inversion pitfalls

Bill and Ted's misadventure.

Bill and Ted are geo-chemists who wish to measure the number of grams of each of three different minerals A, B, C held in a single rock sample.

Let
a be the number of grams of A,
b be the number of grams of B,
c be the number of grams of C
d be the number of grams in the sample.

From Todd Will

By performing complicated experiments Bill and Ted are able to measure four relationships between *a,b,c,d* which they record in the matrix below:

$$\begin{pmatrix} 93.477 & 10.202 & -28.832 \\ 1.93 & 32.816 & 62.414 \\ 26.821 & 36.816 & 57.234 \\ 23.2134 & -86.3925 & 44.693 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 34.7177 \\ 70.9241 \\ 82.9271 \\ -26.222 \end{pmatrix}$$

$$Ax = b$$

Now we have more equations than we need

What to do?

From Todd Will

One thing to do is throw out one of the equations

(in reality only a Mathematician is naïve enough to think that three equations is sufficient to solve for three unknowns – but lets try it anyway).

So throw out one - leaving

$$\begin{pmatrix} 93.477 & 10.202 & -28.832 \\ 1.93 & 32.816 & 62.414 \\ 26.821 & 36.816 & 57.234 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 34.7177 \\ 70.9241 \\ 82.9271 \end{pmatrix}$$

$$Ax = b$$

(different *A* and *b* from before)

From Todd Will

Remembering some of their linear algebra the know that the matrix is not invertible if the determinant is zero, so they check that

$$\begin{vmatrix} 93.477 & 10.202 & -28.832 \\ 1.93 & 32.816 & 62.414 \\ 26.821 & 36.816 & 57.234 \end{vmatrix} \approx -2$$

OK so far

(or "fat, dumb and happy")

From Todd Will

So now we can compute

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 93.477 & 10.202 & -28.832 \\ 1.93 & 32.816 & 62.414 \\ 26.821 & 36.816 & 57.234 \end{pmatrix}^{-1} \begin{pmatrix} 34.7177 \\ 70.9241 \\ 82.9271 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0.7 \end{pmatrix}$$

$$x = A^{-}b$$

So now we're done.

From Todd Will

Or are we?

Next they realize that the measurements are really only good to 0.1

So they round to 0.1 and do it again

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 93.477 & 10.202 & -28.832 \\ 1.93 & 32.816 & 62.414 \\ 26.821 & 36.816 & 57.234 \end{pmatrix}^{-1} \begin{pmatrix} 34.7 \\ 70.9 \\ 82.9 \end{pmatrix} = \begin{pmatrix} -1.68294 \\ 8.92282 \\ -3.50254 \end{pmatrix}$$

$$x = A^{-}b$$

From Todd Will

Now they notice a small problem –

They get a very different answer

(and they don't notice they have a bigger problem that they have negative weights/amounts!)

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0.7 \end{pmatrix} \qquad \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -1.68294 \\ 8.92282 \\ -3.50254 \end{pmatrix}$$

From Todd Will

So what's the problem?

First find the SVD of A.

$$A = \begin{pmatrix} 93.477 & 10.202 & -28.832 \\ 1.93 & 32.816 & 62.414 \\ 26.821 & 36.816 & 57.234 \end{pmatrix} =$$

$$A = \begin{pmatrix} \vec{h}_1 & \vec{h}_2 & \vec{h}_3 \end{pmatrix} \begin{pmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 0.0002 \end{pmatrix} \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vec{a}_3 \end{pmatrix}$$

Since there are three non-zero values on the diagonal A is invertible

From Todd Will

BUT, one of the singular values is much, much less than the others

$$A = \begin{pmatrix} 93.477 & 10.202 & -28.832 \\ 1.93 & 32.816 & 62.414 \\ 26.821 & 36.816 & 57.234 \end{pmatrix} =$$

$$A = \begin{pmatrix} \vec{h}_1 & \vec{h}_2 & \vec{h}_3 \end{pmatrix} \begin{pmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 0.0002 \end{pmatrix} \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vec{a}_3 \end{pmatrix}$$

So the matrix is "almost" rank 2

(which would be non-invertible)

From Todd Will

We can also calculate the SVD of A^{-1}

$$A^{-1} = \left(\vec{a}_1 \; \vec{a}_2 \; \vec{a}_3 \right) \begin{pmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 5000 \end{pmatrix} \begin{pmatrix} \vec{h}_1 \\ \vec{h}_2 \\ \vec{h}_3 \end{pmatrix}$$

From Todd Will

So now we can see what happened

(why the two answers were so different)

Let y be the first version of b

Let y' be the second version of b (to 0.1)

$$\left|A^{-1}y - A^{-1}y'\right| = \left|A^{-1}\left(y - y'\right)\right| =$$

$$\left|\left(\vec{a}_1\ \vec{a}_2\ \vec{a}_3\right)\begin{pmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 5000 \end{pmatrix}\begin{pmatrix} \vec{h}_1 \\ \vec{h}_2 \\ \vec{h}_3 \end{pmatrix}\left(y - y'\right)\right|$$

So A^{-1} stretches vectors parallel to h_3 and a_3 by a factor of 5000.

From Todd Will

Returning to GPS

$$P^{R1}\left(t^R,t^1\right)=\sqrt{\left(x^1\left(t^1\right)-x^R\left(t^R\right)\right)^2+\left(y^1\left(t^1\right)-y^R\left(t^R\right)\right)^2+\left(z^1\left(t^1\right)-z^R\left(t^R\right)\right)^2}+\left(\tau^R-\tau^1\right)c$$

$$P^{R2}\left(t^R,t^2\right)=\sqrt{\left(x^2\left(t^2\right)-x^R\left(t^R\right)\right)^2+\left(y^2\left(t^2\right)-y^R\left(t^R\right)\right)^2+\left(z^2\left(t^2\right)-z^R\left(t^R\right)\right)^2}+\left(\tau^R-\tau^2\right)c$$

$$P^{R3}\left(t^R,t^3\right)=\sqrt{\left(x^3\left(t^3\right)-x^R\left(t^R\right)\right)^2+\left(y^3\left(t^3\right)-y^R\left(t^R\right)\right)^2+\left(z^3\left(t^3\right)-z^R\left(t^R\right)\right)^2}+\left(\tau^R-\tau^3\right)c$$

$$P^{R4}\left(t^R,t^4\right)=\sqrt{\left(x^4\left(t^4\right)-x^R\left(t^R\right)\right)^2+\left(y^4\left(t^4\right)-y^R\left(t^R\right)\right)^2+\left(z^4\left(t^4\right)-z^R\left(t^R\right)\right)^2}+\left(\tau^R-\tau^4\right)c$$

We have 4 unknowns (x^R,y^R,z^R and τ^R)
And 4 (nonlinear) equations
(later we will allow more satellites)
So we can solve for the unknowns

Again, we cannot solve this directly

Will solve interatively by

1) Assuming a location

2) Linearizing the range equations

3) Use least squares to compute new (better) location

4) Go back to 1 using location from 3

We do this till some convergence criteria is met (if we're lucky)

linearize

So - for one satellite we have

$$P_{\text{observed}} = P_{\text{model}} + \vec{v}$$

$$P_{\text{observed}} = P(x, y, z, \tau) + \vec{v}$$

linearize

$$P(x,y,z,\tau) \approx P(x_0,y_0,z_0,\tau_0) + (x - x_0)\frac{\partial P}{\partial x}\bigg|_{(x_0,y_0,z_0,\tau_0)}$$

$$+ (y - y_0)\frac{\partial P}{\partial y}\bigg|_{(x_0,y_0,z_0,\tau_0)} + (z - z_0)\frac{\partial P}{\partial z}\bigg|_{(x_0,y_0,z_0,\tau_0)} + (\tau - \tau_0)\frac{\partial P}{\partial \tau}\bigg|_{(x_0,y_0,z_0,\tau_0)}$$

$$P(x,y,z,\tau) \approx P_{\text{computed}} + \frac{\partial P}{\partial x}\Delta x + \frac{\partial P}{\partial y}\Delta y + \frac{\partial P}{\partial z}\Delta z + \frac{\partial P}{\partial \tau}\Delta \tau$$

Residual

Difference between observed and calculated (linearized)

$$P_{\text{observed}} = P(x, y, z, \tau) + \vec{v}$$

$$\Delta P = P_{\text{observed}} - P_{\text{computed}}$$

$$\Delta P = P(x, y, z, \tau) + \vec{v} - P_{\text{computed}}$$

$$\Delta P = P_{\text{computed}} + \frac{\partial P}{\partial x}\Delta x + \frac{\partial P}{\partial y}\Delta y + \frac{\partial P}{\partial z}\Delta z + \frac{\partial P}{\partial \tau}\Delta \tau + \vec{v} - P_{\text{computed}}$$

$$\Delta P = \frac{\partial P}{\partial x}\Delta x + \frac{\partial P}{\partial y}\Delta y + \frac{\partial P}{\partial z}\Delta z + \frac{\partial P}{\partial \tau}\Delta \tau + \vec{v}$$

Blewitt, Basics of GPS in "Geodetic Applications of GPS"

So we have the following for one satellite

$$\Delta P = \frac{\partial P}{\partial x}\Delta x + \frac{\partial P}{\partial y}\Delta y + \frac{\partial P}{\partial z}\Delta z + \frac{\partial P}{\partial \tau}\Delta \tau + \vec{v}$$

Which we can recast in matrix form

$$\Delta P = \begin{pmatrix} \frac{\partial P}{\partial x} & \frac{\partial P}{\partial y} & \frac{\partial P}{\partial z} & \frac{\partial P}{\partial \tau} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta \tau \end{pmatrix} + \vec{v}$$

For m satellites (where $m \geq 4$)

$$\begin{pmatrix} \Delta P^1 \\ \Delta P^1 \\ \Delta P^1 \\ \vdots \\ \Delta P^m \end{pmatrix} = \begin{pmatrix} \frac{\partial P^1}{\partial x} & \frac{\partial P^1}{\partial y} & \frac{\partial P^1}{\partial z} & \frac{\partial P^1}{\partial \tau} \\ \frac{\partial P^2}{\partial x} & \frac{\partial P^2}{\partial y} & \frac{\partial P^2}{\partial z} & \frac{\partial P^2}{\partial \tau} \\ \frac{\partial P^3}{\partial x} & \frac{\partial P^3}{\partial y} & \frac{\partial P^3}{\partial z} & \frac{\partial P^3}{\partial \tau} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial P^m}{\partial x} & \frac{\partial P^m}{\partial y} & \frac{\partial P^m}{\partial z} & \frac{\partial P^m}{\partial \tau} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta \tau \end{pmatrix} + \begin{pmatrix} \nu^1 \\ \nu^2 \\ \nu^3 \\ \vdots \\ \nu^n \end{pmatrix}$$

Which is usually written as

$$\vec{b} = A\vec{x} + \vec{v}$$

Calculate the derivatives

$$\frac{\partial P^{RS}}{\partial x^R} = \sqrt{\left(x^S\left(t^S\right) - x^R\left(t^R\right)\right)^2 + \left(y^S\left(t^S\right) - y^R\left(t^R\right)\right)^2 + \left(z^S\left(t^S\right) - z^R\left(t^R\right)\right)^2} + \left(\tau^R - \tau^1\right) c$$

$$\frac{\partial P^{RS}}{\partial x^R} = \frac{\left((-1)\left(\frac{1}{2}\right) 2\right)\left(x^S\left(t^S\right) - x^R\left(t^R\right)\right)}{\sqrt{\left(x^S\left(t^S\right) - x^R\left(t^R\right)\right)^2 + \left(y^S\left(t^S\right) - y^R\left(t^R\right)\right)^2 + \left(z^S\left(t^S\right) - z^R\left(t^R\right)\right)^2}}$$

$$\frac{\partial P^{RS}}{\partial x^R} = \frac{\left(x^R\left(t^R\right) - x^S\left(t^S\right)\right)}{\rho^R}, \text{ similarly for } y \text{ and } z$$

$$\frac{\partial P^{RS}}{\partial \tau^R} = c$$

Blewitt, Basics of GPS in "Geodetic Applications of GPS"

So we get

$$A = \begin{pmatrix} \dfrac{x_0 - x^1}{\rho^1} & \dfrac{y_0 - y^1}{\rho^1} & \dfrac{z_0 - z^1}{\rho^1} & c \\ \dfrac{x_0 - x^2}{\rho_2} & \dfrac{y_0 - y^2}{\rho_2} & \dfrac{z_0 - z^2}{\rho_2} & c \\ \dfrac{x_0 - x^3}{\rho_3} & \dfrac{y_0 - y^3}{\rho_3} & \dfrac{z_0 - z^3}{\rho_3} & c \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{x_0 - x^m}{\rho_m} & \dfrac{y_0 - y^m}{\rho_m} & \dfrac{z_0 - z^m}{\rho_m} & c \end{pmatrix}$$

Is function of direction to satellite

Note last column is a constant

Blewitt, Basics of GPS in "Geodetic Applications of GPS"

Consider some candidate solution *x'*

Then we can write

$$\vec{v} = b - \boxed{A\vec{x}'}$$

b are the observations

v hat are the residuals

We would like to find the *x'* that minimizes the *v hat*

So the question now is how to find this x'

One way, and the way we will do it,

Least Squares

Since we have already done this – we'll go fast

Use solution to linearized form of observation equations to write estimated residuals

$$\hat{v} = \vec{b} - A\hat{x}'$$

Vary value of x to minimize

$$J(\vec{x}) = \sum_{i=1}^{m} \vec{v}_i^2 = \vec{v}^T \vec{v} = \left(\vec{b} - A\vec{x}\right)^T \left(\vec{b} - A\vec{x}\right)$$

$$\delta J(\hat{x}) = 0$$

$$\delta\left\{\left(\vec{b} - A\hat{x}\right)^T\left(\vec{b} - A\hat{x}\right)\right\} = 0$$

$$\left\{\delta\left(\vec{b} - A\hat{x}\right)^T\right\}\left(\vec{b} - A\hat{x}\right) + \left(\vec{b} - A\hat{x}\right)^T\left\{\delta\left(\vec{b} - A\hat{x}\right)\right\} = 0$$

$$\left(-A\delta\,\hat{x}\right)^T\left(\vec{b} - A\hat{x}\right) + \left(\vec{b} - A\hat{x}\right)^T\left(-A\delta\,\hat{x}\right) = 0$$

$$\left(-A\delta\,\hat{x}\right)^T\left(\vec{b} - A\hat{x}\right) + \left(\vec{b} - A\hat{x}\right)^T\left(-A\delta\,\hat{x}\right) = 0$$

$$\left(A\delta\,\hat{x}\right)^T\left(\vec{b} - A\hat{x}\right) = 0$$

$$\left(\delta\,\hat{x}^T A^T\right)\left(\vec{b} - A\hat{x}\right) = 0$$

$$\delta\,\hat{x}^T\left(A^T\vec{b} - A^T A\hat{x}\right) = 0$$

$$\boxed{A^T\vec{b} = A^T A\hat{x} \qquad \text{Normal equations}}$$

$$\hat{x} = \left(A^T A\right)^{-1} A^T\vec{b} \qquad \text{Solution to normal equations}$$

$$\hat{x} = \left(A^T A\right)^{-1} A^T \vec{b}$$

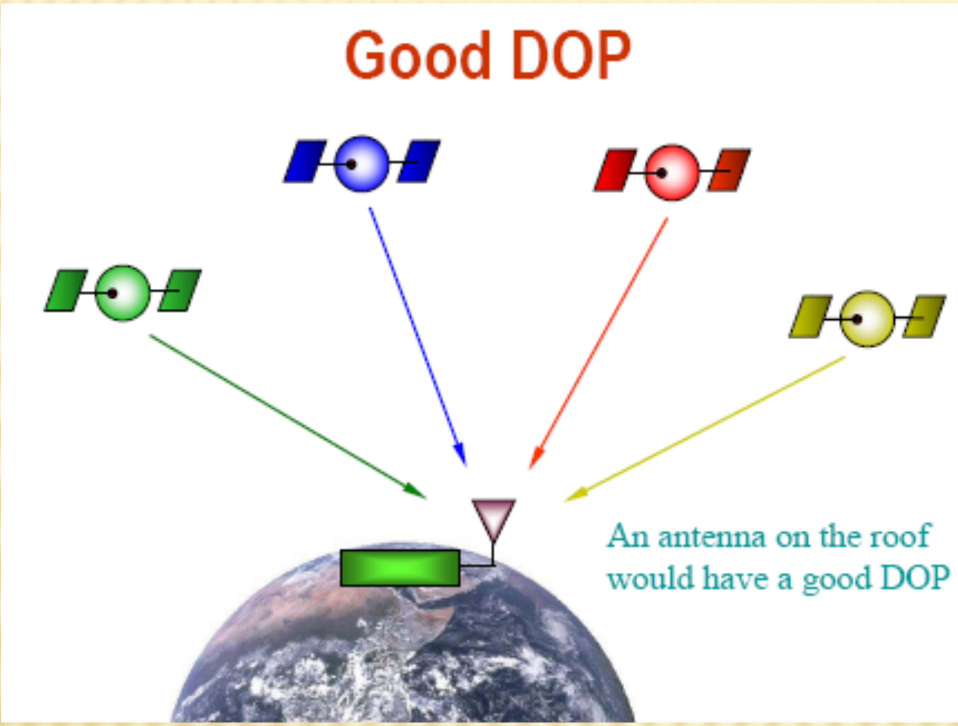
Assumes

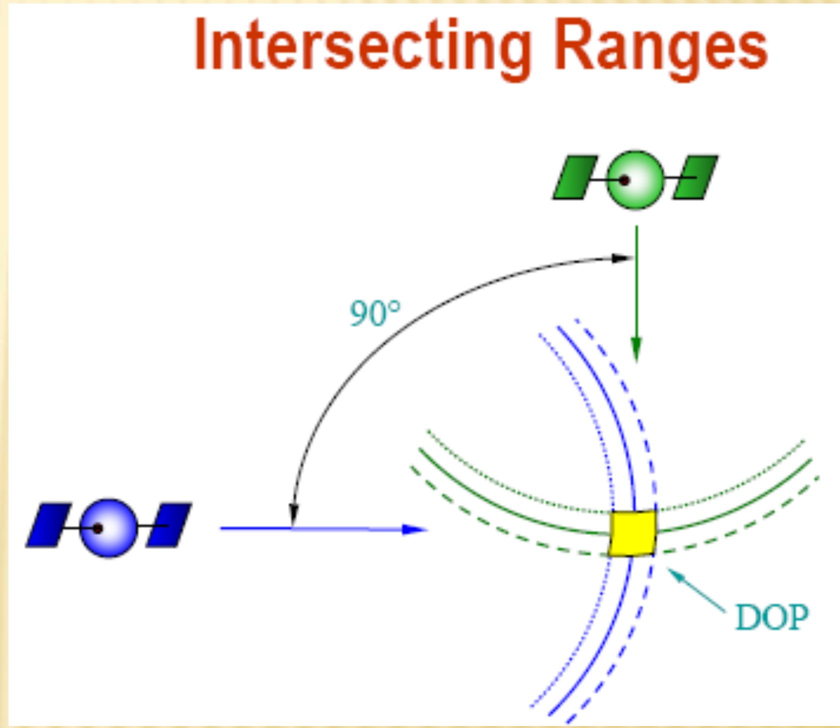
Inverse exists

(m greater than or equal to 4, necessary but not sufficient condition)

Can have problems similar to earthquake locating (two satellites in “same” direction for example – has effect of reducing rank by one)

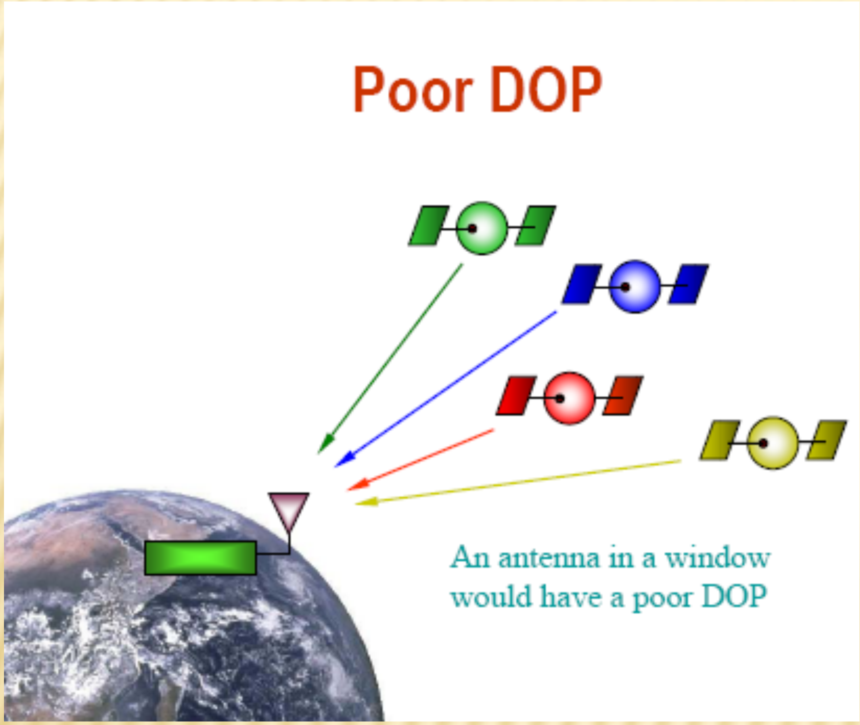
GPS tutorial Signals and Data

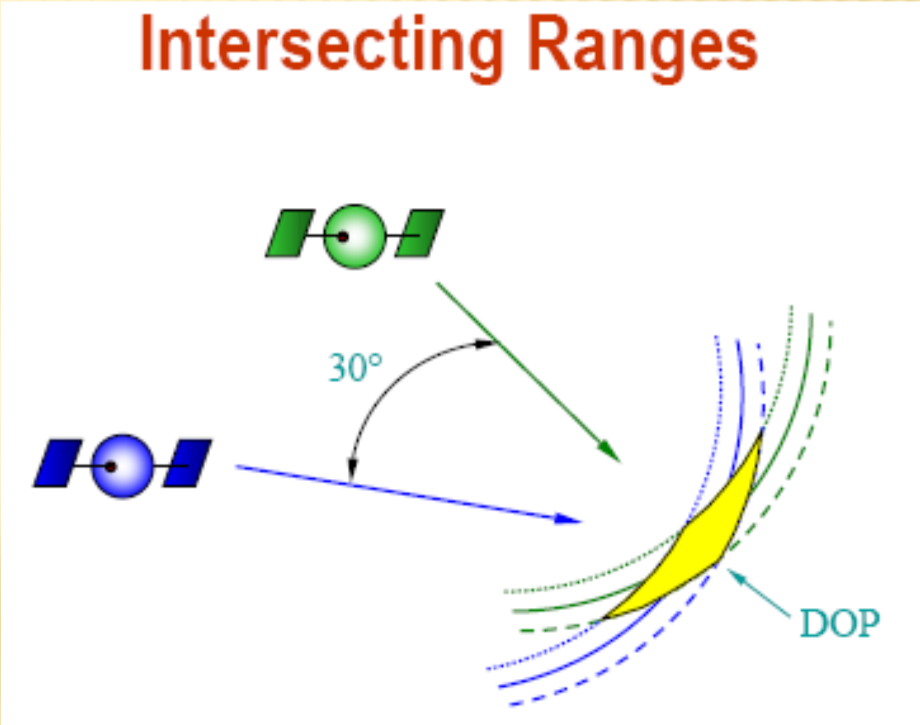




http://www.unav-micro.com/about_gps.htm

GPS tutorial Signals and Data





http://www.unav-micro.com/about_gps.htm

Elementary Concepts

Variables:

things that we measure, control, or manipulate in research. They differ in many respects, most notably in the role they are given in our research and in the type of measures that can be applied to them.

Observational vs. experimental research.

Most empirical research belongs clearly to one of those two general categories.

In observational research we do not (or at least try not to) influence any variables but only measure them and look for relations (correlations) between some set of variables.

In experimental research, we manipulate some variables and then measure the effects of this manipulation on other variables.

From G. Mattioli

Observational vs. experimental research.

Dependent vs. independent variables.

Independent variables are those that are manipulated whereas dependent variables are only measured or registered.

From G. Mattioli

Variable Types and Information Content

Measurement scales.

Variables differ in "how well" they can be measured.

Measurement error involved in every measurement, which determines the "amount of information" obtained.

Another factor is the variable's "type of measurement scale."

From G. Mattioli

Variable Types and Information Content

Nominal variables

allow for only qualitative classification.

That is, they can be measured only in terms of whether the individual items belong to some distinctively different categories, but we cannot quantify or even rank order those categories.

Typical examples of nominal variables are gender, race, color, city, etc.

From G. Mattioli

Variable Types and Information Content

Ordinal variables

allow us to rank order the items we measure

in terms of which has less and which has more of the quality represented by the variable, but still they do not allow us to say "how much more."

A typical example of an ordinal variable is the socioeconomic status of families.

From G. Mattioli

Variable Types and Information Content

Interval variables

allow us not only to rank order the items that are measured,

but also to quantify and compare the sizes of differences between them.

For example, temperature, as measured in degrees Fahrenheit or Celsius, constitutes an interval scale.

From G. Mattioli

Variable Types and Information Content

Ratio variables

are very similar to interval variables;

in addition to all the properties of interval variables,

they feature an identifiable absolute zero point, thus they allow for statements such as x is two times more than y.

Typical examples of ratio scales are measures of time or space.

Systematic and Random Errors

Error:

Defined as the difference between a calculated or observed value and the "true" value

Systematic and Random Errors

Blunders:

Usually apparent

either as obviously incorrect data points or results that are not reasonably close to the expected value.

Easy to detect (usually).

Easy to fix (throw out data).

From G. Mattioli

Systematic and Random Errors

Systematic Errors:

Errors that occur reproducibly from faulty calibration of equipment or observer bias.

Statistical analysis in generally not useful,

but rather corrections must be made based on experimental conditions.

From G. Mattioli

Systematic and Random Errors

Random Errors:

Errors that result from the fluctuations in observations.

Requires that experiments be repeated a sufficient number of time to establish the precision of measurement.

(statistics useful here)

From G. Mattioli

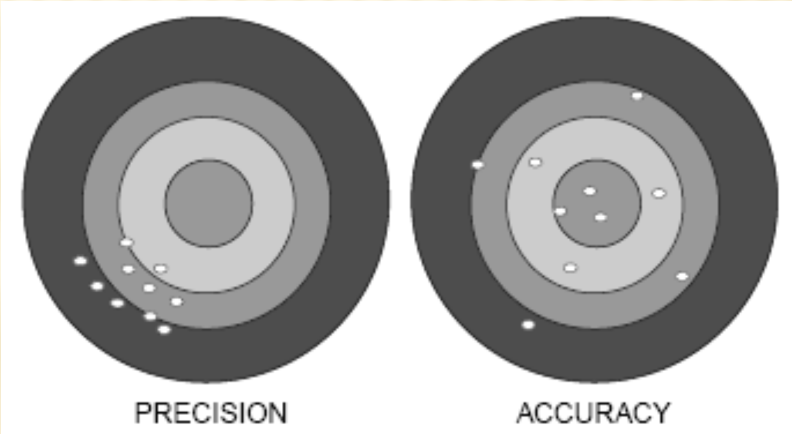
Accuracy vs. Precision

From G. Mattioli

Accuracy vs. Precision

Accuracy: A measure of how close an experimental result is to the true value.

Precision: A measure of how exactly the result is determined. It is also a measure of how reproducible the result is.



From G. Mattioli

Accuracy vs. Precision

Absolute precision:

indicates the uncertainty in the same units as the observation

Relative precision:

indicates the uncertainty in terms of a fraction of the value of the result

From G. Mattioli

Uncertainties

In most cases,

cannot know what the "true" value is unless there is an independent determination

(*i.e.* different measurement technique).

Uncertainties

Only can consider estimates of the error.

Discrepancy is the difference between two or more observations. This gives rise to uncertainty.

Probable Error:
Indicates the magnitude of the error we estimate to have made in the measurements.
Means that if we make a measurement that we will be wrong by that amount "on average".

From G. Mattioli

Parent vs. Sample Populations

Parent population:

Hypothetical probability distribution if we were to make an infinite number of measurements of some variable or set of variables.

Parent vs. Sample Populations

Sample population:

Actual set of experimental observations or measurements of some variable or set of variables.

In General:

(Parent Parameter) = limit (Sample Parameter)

When the number of observations, N, goes to infinity.

From G. Mattioli

some univariate statistical terms:

mode:

value that occurs most frequently in a distribution

(usually the highest point of curve)

may have more than one mode
(eg. Bimodal – example later)
in a dataset

From G. Mattioli

some univariate statistical terms:

median:

value midway in the frequency distribution
…half the area under the curve is to right and other to left

mean:

arithmetic average
…sum of all observations divided by # of observations

the mean is a poor measure of central tendency in skewed distributions

From G. Mattioli

Average, mean or expected value for random variable

$$E(x) = \mu = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} x_i$$

(more general) if have probability for each x_i

$$E(x) = \mu = \lim_{n \to \infty} \sum_{i=1}^{n} p_i x_i$$

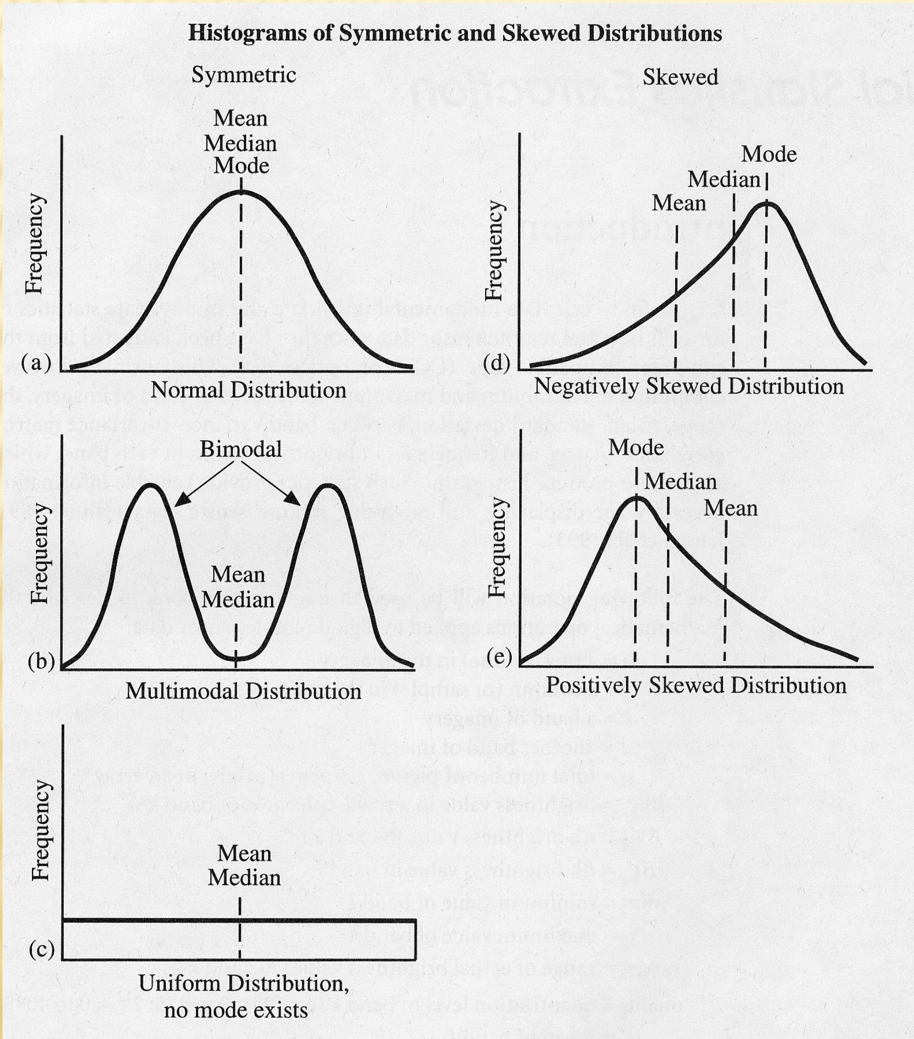
some univariate statistical terms:

range: measure of dispersion about mean

(maximum minus minimum)

when max and min are unusual values, range may be a misleading measure of dispersion

Histogram
useful graphic representation of information content of sample or parent population



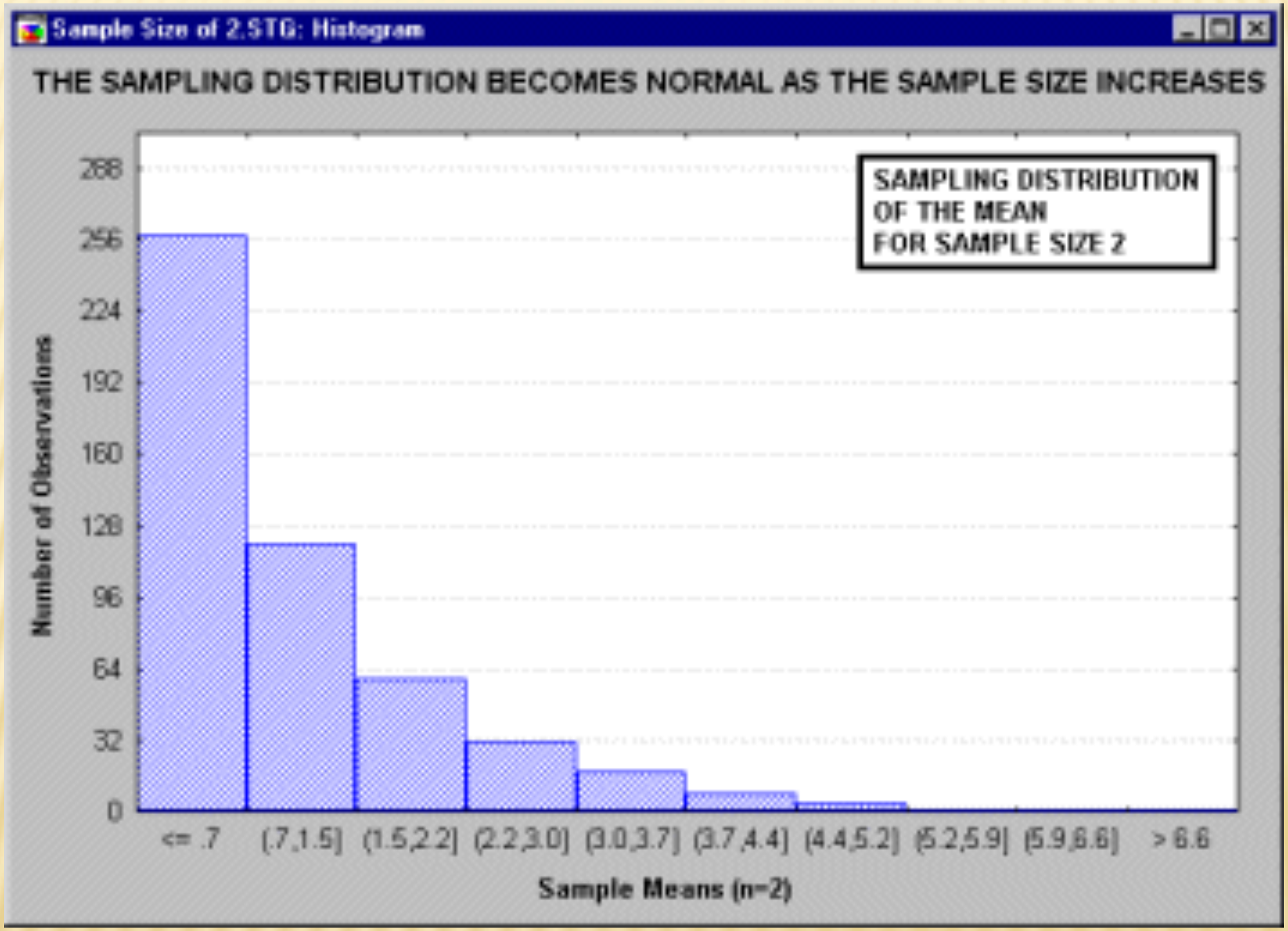
many statistical tests assume values are normally distributed

not always the case! examine data prior to processing

from: Jensen, 1996

From G. Mattioli

Distribution vs. Sample Size



http://dhm.mstu.edu.ru/e_library/statistica/textbook/graphics/

Deviations

The deviation, δ_i, of any measurement x_i from the mean *m* of the parent distribution is defined as the difference between x_i and m

$$\delta x_i = x_i - \mu$$

From G. Mattioli

Deviations

Average deviation, α,

is defined as the average of the magnitudes of the deviations,

Magnitudes given by the absolute value of the deviations.

$$\alpha = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left| x_i - \mu \right|$$

From G. Mattioli

Root mean square

$$RMS = \lim_{n \to \infty} \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2}$$

Of deviations or residuals – standard deviation

$$\sigma = \lim_{n \to \infty} \sqrt{\frac{1}{n} \sum_{i=1}^{n} \delta_i^2}$$

Sample Mean and Standard Deviation

For a series of *n* observations, the most probable estimate of the mean μ is the average of the observations.

We refer to this as the sample mean to distinguish it from the parent mean μ.

$$\mu \approx \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

From G. Mattioli

Sample Mean and Standard Deviation

Our best estimate of the standard deviation σ would be from:

$$\sigma^2 = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\mu\right)^2 \approx \frac{1}{n}\sum_{i=1}^{n}\left(x_i-\mu\right)^2$$

But we cannot know the true parent mean μ so the best estimate of the sample variance and standard deviation would be:

$$\sigma^2 = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$$ Sample Variance

From G. Mattioli

Some other forms to write variance

$$\sigma^2 = VAR(x) = \frac{1}{(n-1)}\sum_{i=1}^{n}\left(x_i - E(x)\right)^2 = \frac{1}{(n-1)}\left(\sum_{i=1}^{n} x_i^{\,2} - N\bar{x}\right)$$

$$\delta x_i = x_i - \bar{x} \qquad\qquad \sigma^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}\delta x_i^{\,2}$$

If have probability for each x_i

$$\sigma^2 = VAR(x) = \sum_{i=1}^{n} p_i\left(x_i - E(x)\right)^2$$

The standard deviation

$$\sigma = \sqrt{VAR(x)} = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^{n} \delta x_i^{\,2}}$$

(Normalization decreased from *N* to (*N* – 1) for the "sample" variance, as μ is used in the calculation)

For a scalar random variable or measurement with a Normal (Gaussian) distribution,

the probability of being within one σ of the mean is 68.3%

small std dev:

observations are clustered tightly about the mean

large std dev:

observations are scattered widely about the mean

Distributions

Binomial Distribution: Allows us to define the probability, *p*, of observing *x* a specific combination of *n* items, which is derived from the fundamental formulas for the permutations and combinations.

Permutations: Enumerate the number of permutations, *Pm(n,x)*, of coin flips, when we pick up the coins one at a time from a collection of *n* coins and put *x* of them into the "heads" box.

$$P_m(n,x) = \frac{n!}{(n-x)!}$$

From G. Mattioli

Combinations:

Relates to the number of ways we can combine the various permutations enumerated above from our coin flip experiment.

Thus the number of combinations is equal to the number of permutations divided by the degeneracy factor *x!* of the permutations (number indistinguishable permutations) .

$$C(n,x) = \frac{P_m(n,x)}{x!} = \frac{n!}{x!(n-x)!} = \binom{n}{x}$$

From G. Mattioli

Probability and the Binomial Distribution

Coin Toss Experiment: If p is the probability of success (landing heads up) is not necessarily equal to the probability $q = 1 - p$ for failure (landing tails up) because the coins may be lopsided!

The probability for each of the combinations of x coins heads up and $n - x$ coins tails up is equal to $p^x q^{n-x}$.

The binomial distribution can be used to calculate the probability:

From G. Mattioli

Probability and the Binomial Distribution

The binomial distribution can be used to calculate the probability of x "successes
in n tries where the individual probabliliyt is p:

$$P_B(n,x,p) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

The coefficients $P_B(x,n,p)$ are closely related to the binomial theorem for the expansion of a power of a sum

$$(p+q)^n = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x}$$

From G. Mattioli

Mean and Variance: Binomial Distribution

$$\mu = \left[\sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^{x} (1-p)^{n-x} \right] = np$$

The average of the number of successes will approach a mean value μ

given by the probability for success of each item *p* times the number of items.

For the coin toss experiment *p=1/2*, half the coins should land heads up on average.

From G. Mattioli

Mean and Variance: Binomial Distribution

The standard deviation is

$$\sigma^2 = \sum_{x=0}^{n}\left[(x-\mu)^2 \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}\right] = np(1-p)$$

If the the probability for a single success *p* is equal to the probability for failure *p=q=1/2*,

the final distribution is symmetric about the mean,

and mode and median equal the mean.

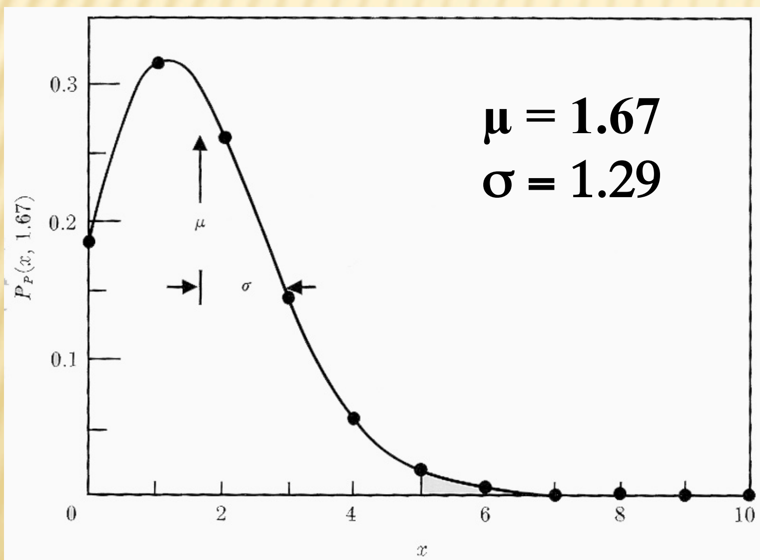
The variance, $\sigma^2 = m/2$.

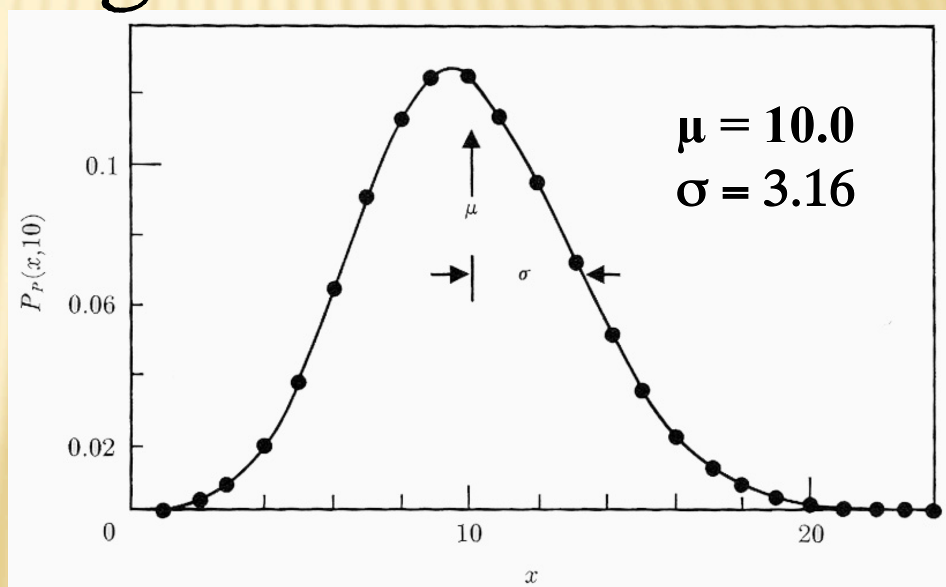
From G. Mattioli

Other Probability Distributions: Special Cases

Poisson Distribution: Approximation to binomial distribution for special case when average number of successes is very much smaller than possible number *i.e.* $\mu \ll n$ because $p \ll 1$.

Distribution is NOT necessarily symmetric! Data are usually bounded on one side and not the other. Advantage $\sigma^2 = m$.



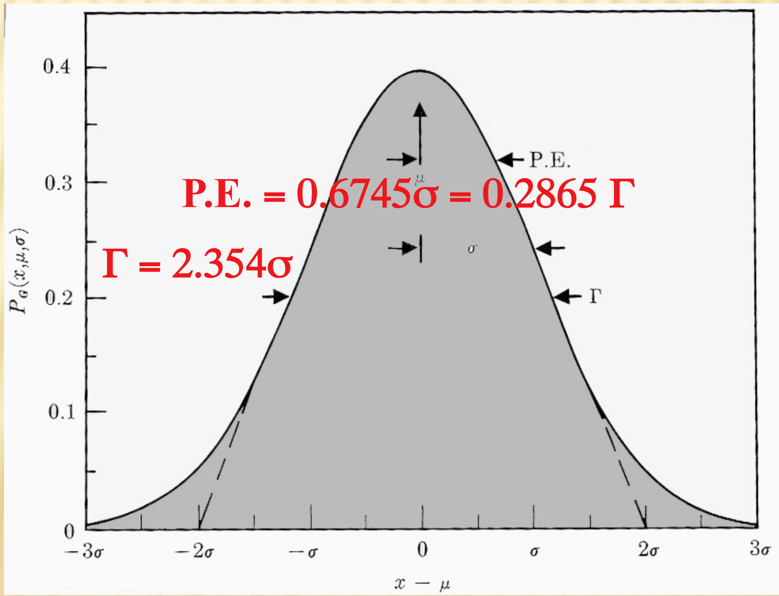


From G. Mattioli

Gaussian or Normal Error Distribution

Gaussian Distribution: Most important probability distribution in the statistical analysis of experimental data.

Functional form is relatively simple and the resultant distribution is reasonable.

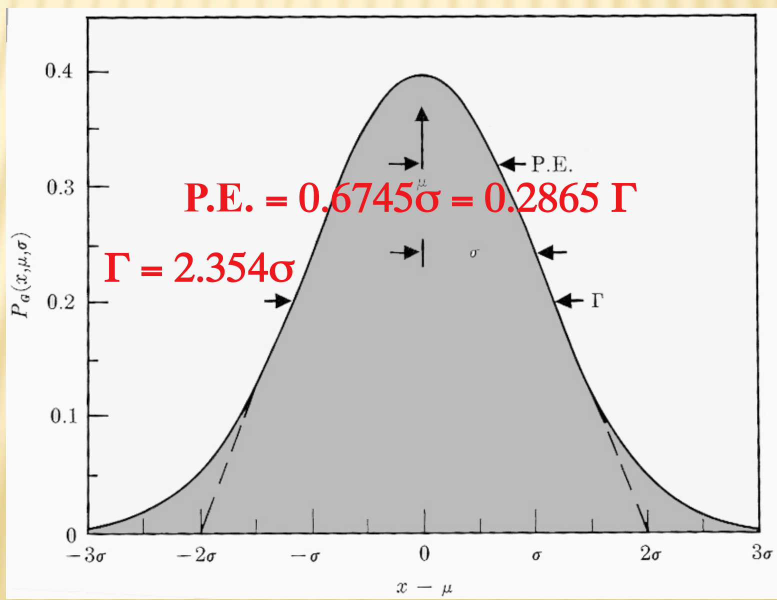


From G. Mattioli

Gaussian or Normal Error Distribution

Another special limiting case of binomial distribution where the number of possible different observations, *n*, becomes infinitely large yielding *np* >> 1.

Most probable estimate of the mean μ from a random sample of observations is the average of those observations!

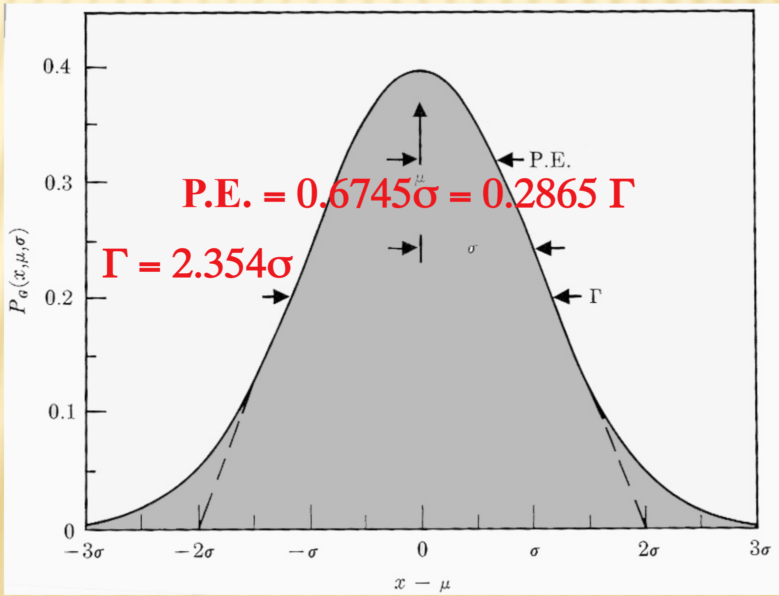


Gaussian or Normal Error Distribution

Probable Error (P.E.) is defined as the absolute value of the deviation such that P_G of the deviation of any random observation is < ½

Tangent along the steepest portionof the probability curve intersects at $e^{-1/2}$ and intersects x axis at the points

$x \approx \mu \pm 2\sigma$



From G. Mattioli

For gaussian / normal error distributions:

Total area underneath curve is 1.00 (100%)

68.27% of observations lie within ± 1 std dev of mean
95% of observations lie within ± 2 std dev of mean
99% of observations lie within ± 3 std dev of mean

Variance, standard deviation, probable error, mean, and weighted root mean square error are commonly used statistical terms in geodesy.

compare (rather than attach significance to numerical value)

From G. Mattioli