Understanding Least Squares. What it is. How to use it.

Charles “Chuck” Ghilani, Ph.D.
Professor Emeritus in Surveying Engineering
Penn State University

Who Am I?

- Received a BS-ED degree in Mathematics and Education in 1974 from UW-Milwaukee
- Taught 6 years at various levels of grade and high school, Asst Coach in football and baseball going to state championship game in football
- Coached many undefeated seasons in girl's little and pony league
  - Coached several years in women's league with several championships – Met Mary (wife) there – one little mistake
- Received a Ph.D. in Civil Engineering from UW-Madison in 1989
- Started B.S. Surveying Engineering program at PSU in 1994
- Past president of *Surveying and Geomatics Educators Society* (SaGES) and *American Association for Geodetic Surveying* (AAGS)
- Past editor of *Surveying and Land Information Science*
- Received the following awards/recognitions in no particular order
  - Earle J. Fennell Award
  - AAGS Fellow
  - PSLS Founders Award
  - PSLS Honorary Member
  - SaGES Educators Award
  - Joseph Dracup Lifetime Achievement Award
- FARM BOY FROM WISCONSIN! – Spent 38 years working on family dairy & hog farm

Class Etiquette

- Turn off all cell phones
  - Or set them to vibrate
  - Leave the room to answer any phone calls
- Ask questions at any point during the class.
  - Simply speak up so that all can hear your question

Course Description

This workshop will cover

1. Basic concepts/mathematics behind least squares adjustments
2. Explore methods used to analyze data after an adjustment
3. Look at post-adjustment tools such as residual analysis and the tau criterion to uncover blunders in data
4. Look at how error ellipses, radial errors, and root mean squares errors are computed and used

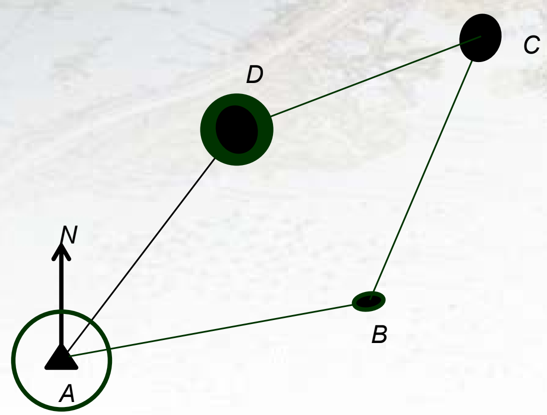
Error Propagation in Traverse Computations

Errors propagate

- Latitudes and departures from angle and distance observations
  - Computation of area/distance/directions from coordinates

Traverse Example

And then we adjust the traverse, which drives errors away from the control



First course
- Uncertainty in distance and azimuth

Second and following courses
- Uncertainty in distance, angle, and coordinates

These errors then propagate to
coordinates for stations,
adjusted distances and azimuths,
and areas

Thus, the largest error is <u>always</u> away the control station
to the station farthest (in connections) from the control station

Discovering Errors in a Horizontal Survey

- For an optical survey
  - Indicators that the observations contain errors
    - Sum of the angles in a polygon traverse do not add to $(n - 2)180°$
    - Sum of the angles in the horizon do not sum to 360°
    - Sum of the latitudes(departures) do not sum to 0
    - Sum of the elevation differences in a closed loop do not sum to 0
- Question: What are your indicators for a GNSS survey?
  - Recognize that the values you get from a GNSS survey are already the results of several least squares adjustments. Even in a RTK and RTN survey

Importance of Adjustments

- Size of errors can be assessed enabling valid decisions to be made regarding acceptance or rejection of observations
- All quantities in a survey are mathematically consistent
- Precisions of final quantities are increased

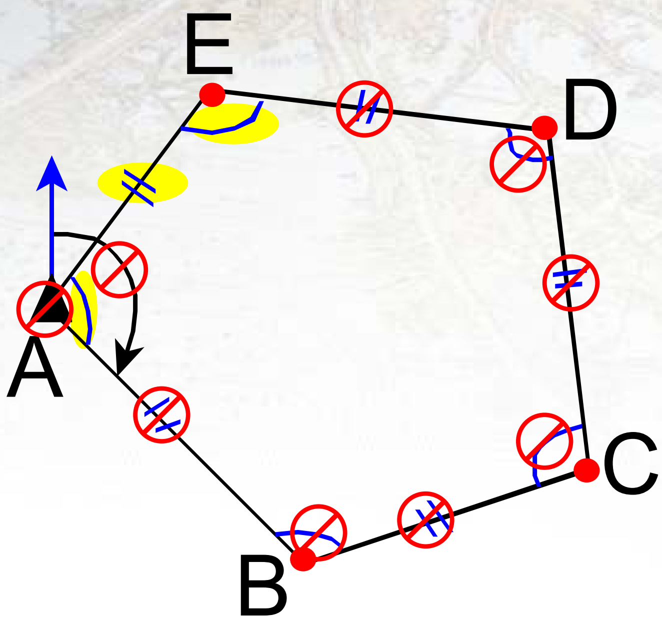
Redundant Observations

- Definition
  - Number of observations in excess of those absolutely needed to determine the unknown or unknowns
    - Also called "Degrees of Freedom"

- Value
  - They enable discrepancies to be computed
  - They help isolate errors
  - They are the basis of adjustments

How Many Degrees of Freedom?

$(N, E)_A$ and Az_{AB} are given



What observations required to compute

B?

C?

D?

E?

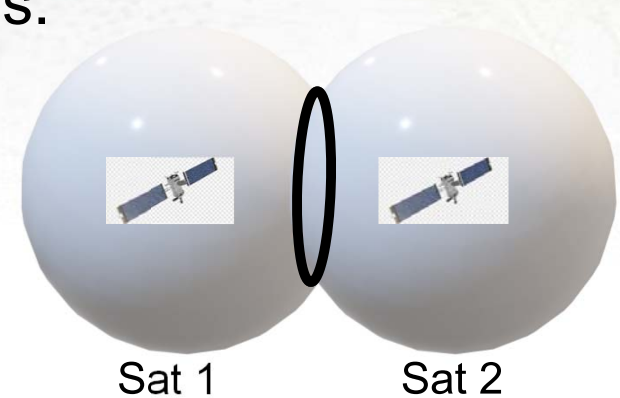
Redundant observations?

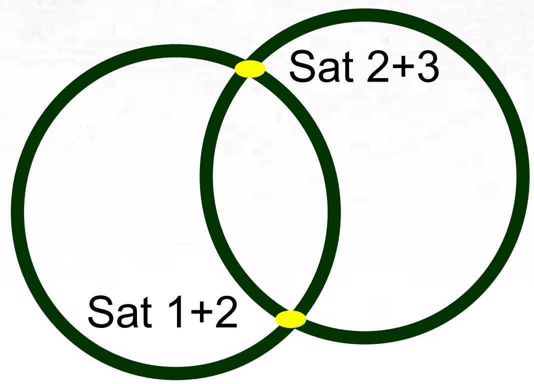
3

Degrees of freedom are also known as redundancies.

GNSS

- To get position from a GNSS receiver you need 4 simultaneously observed satellites for single epoch.
  - 3 satellites determine two possible locations for the receiver
  - The fourth satellite is used to determine the receiver clock bias.



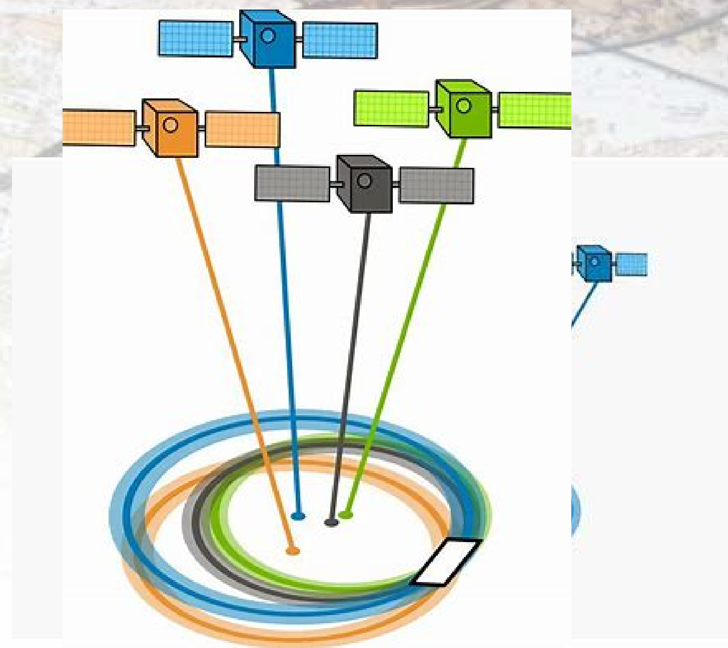


GNSS Positioning

- The accuracy of your position is dependent on
  - Quality of ephemeris
  - Geometry of satellites
  - Atmospheric conditions at time of observation
  - Ionospheric conditions at time of observation
  - Clock biases
  - Etc.

Uncertainty in GNSS Positions

- Depending on geometry of satellites at time of observation, uncertainty in position maybe:
  - Low (good)
  - High (bad)
    - Same as setting two ties stations that are inline with point they are to locate

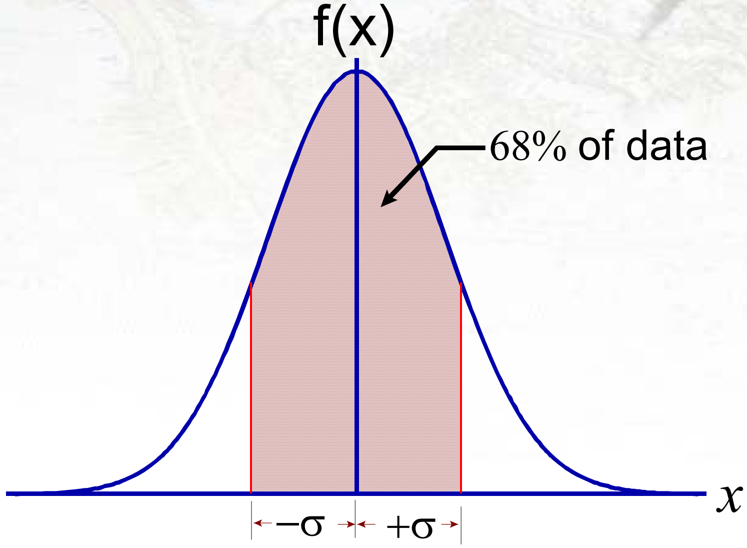


GNSS Observations

- So what if there are more than 4 satellites?
  - Each additional satellite provides one more pseudorange, and thus one more sphere
  - Thus, there are typically a multitude of possible positions for the receiver for each epoch but only one common position!
  - Every epoch provides a similar number of additional observations
  - Least squares is used to find the <u>most probable position</u> of the receiver

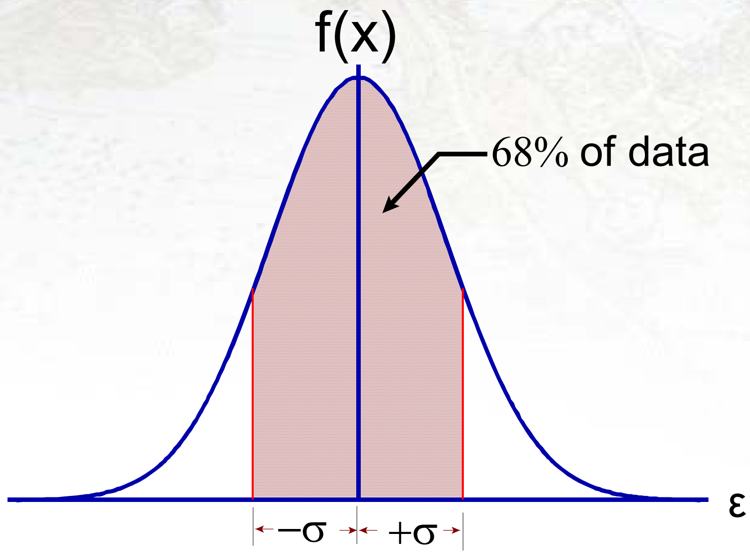
Normal Distribution Curve

- All surveying observations follow a normal distribution curve
  - So let's study the curve!



Normal Distribution Curve

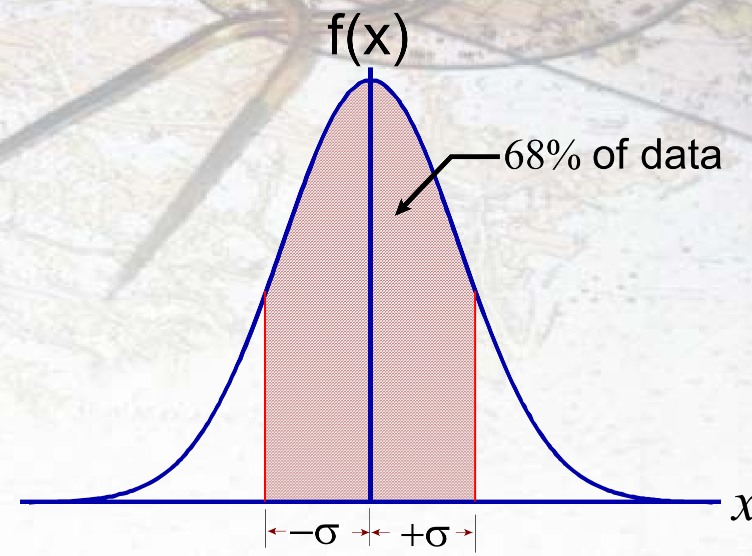
- Plotting errors in observations from a **<u>population</u>** of data
- Normal distribution curve defined by



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\varepsilon^2/2\sigma^2}$$

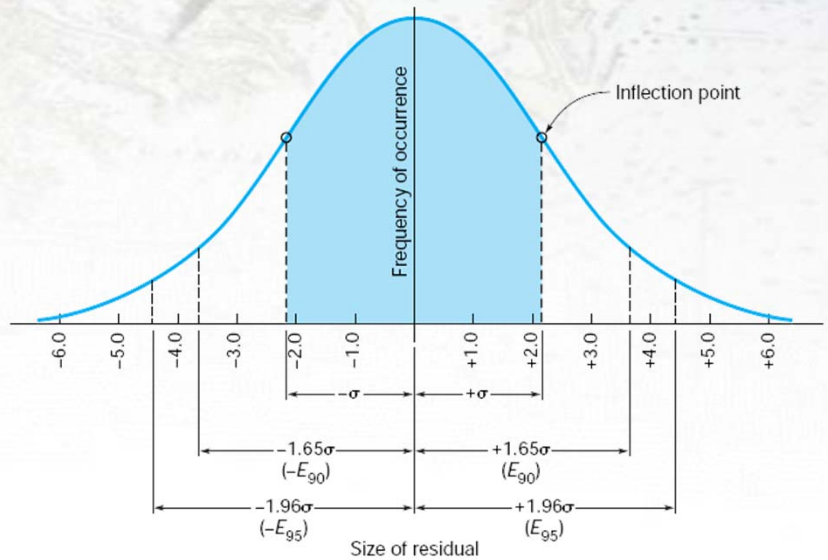
Normal Distribution Curve

- A small standard deviation, σ, indicates a precise set of observations



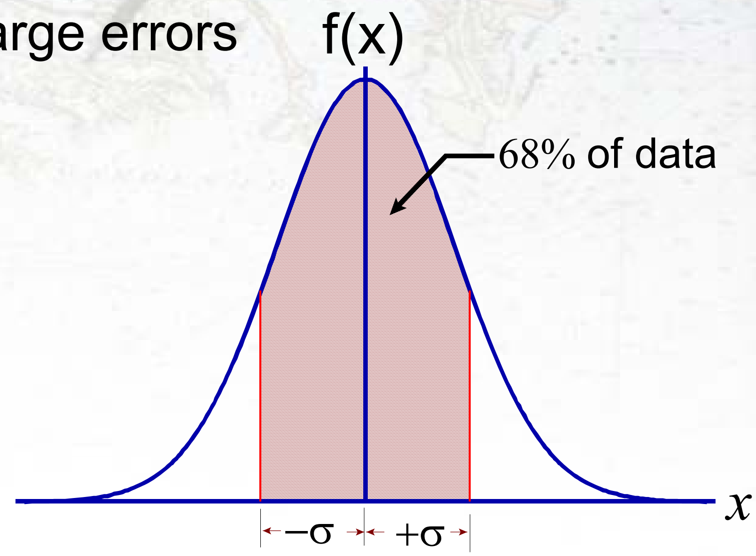
Normal Distribution Curve

- A small standard deviation, *σ*, indicates a precise set of observations
- A large standard deviation, *σ,* indicates a less precise set of observations



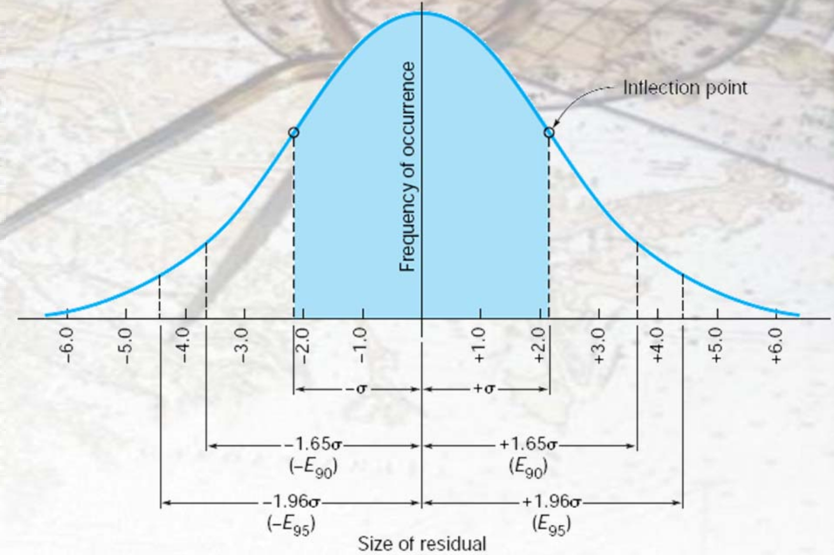
Properties of Distribution

- Some Laws of Probability
  - Positive and negative errors occur with equal probability and frequency
    - If a preponderance of residuals are one sign then there probably is a blunder causing this
  - Small errors occur more frequently than large errors
    - Large errors seldom occur



Useful Probabilities

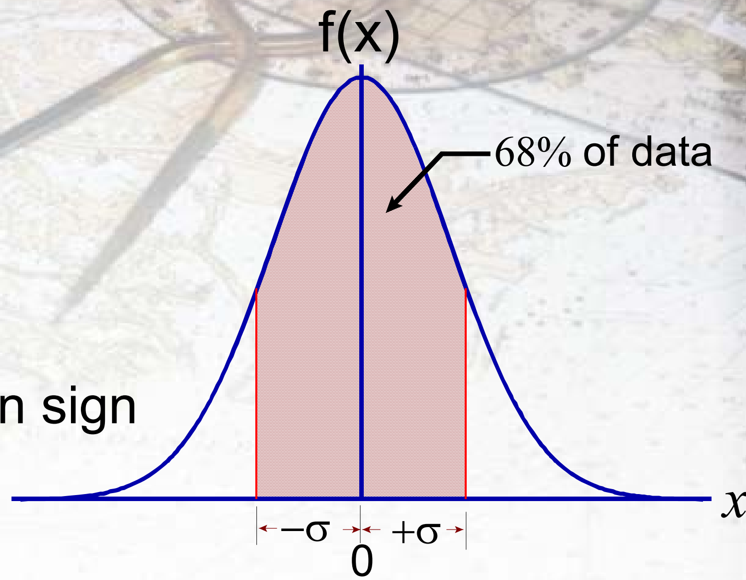
Symbol	Multiplier	Percentage
E_{50}	0.675 σ	50
E_{90}	1.645 σ	90
E_{95}	1.960 σ	95
E_{99}	2.576 σ	99
$E_{99.7}$	2.965 σ	99.7



- Note: For a normally distributed population, 99.7% of data should be within $\pm 2.965\sigma$ ($\sim 3\sigma$) of the mean
- Use these values to define “large” error
  - We will come back to this later!

Properties of Random Errors

- Random errors are
  1. Generally small in magnitude
     - Large random errors seldom occur
  2. Follow the laws of probability
     - Are as likely to be negative as positive in sign
  3. Impossible to avoid



- Use these principles to help identify blunders

What is Least Squares Adjustment?

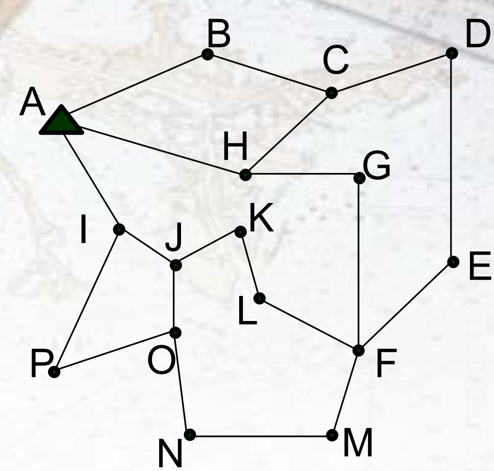
- It is the simultaneous solution of a set of equations for a set of unknown parameters, typically the coordinates and/or heights of stations
- Our observations form the set of equations
  - E.g. $Dist_{AB} + \boldsymbol{v_{AB}} = \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2}$
  - GPS: $\Phi_{L1}^{j}(t) + \boldsymbol{v_{\Phi}} = \frac{1}{\lambda_{L1}}\rho^{j}(t) + f_{L1}\Delta\delta^{j} + N_{L1} - f_{L1}\delta^{Iono} + f_{L1}\delta^{Trop}$
- For every observation there is one equation
  - I.e. Angles, distances, azimuths, carrier frequencies L1, L2, & L5

Least Squares Adjustment

- Developed in the late 1700's by Carl F. Gauss
- Observations adjusted to conform to the laws of probability for a normal distribution
  - For any given data set
    - Provides most probable solution
    - Adjusted observations satisfy geometric constraints
    - There is a single solution
- Observations can be individually weighted
- Adjusts all the observations simultaneously

Advantages of Least Squares

- Errors adjusted according to laws of probability
- How many different solutions in figure to right?
- Lsq. adj. provides a single solution to a set of observations, which is the <u>most probable solution</u>
- Forces observations to satisfy geometric closures
- Can perform post-adjustment analysis
- Can perform presurvey planning



Interesting Fact

- Compass Rule:
  - Developed as response to a contest in the *Mathematical Analyst* in Philadelphia in 1810
    - Bowditch won the contest!
  - Adjustment performed assuming that angles and distances are of equal quality

Interesting Fact

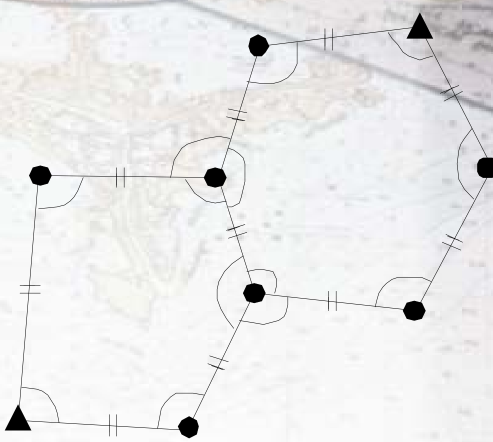
- Compass Rule:
  - So
    - Are all angles and distances of equal quality?
      - Distances? (3 mm + 3 ppm), (5 mm + 5 ppm), taped?
      - Angles? Longer sight distances = better sights on target and smaller uncertainties in angles observed
    - Why are angles adjusted twice?
    - Why are all angles adjusted the same?

Interesting Facts

- Least squares
  - Method developed between 1774 and 1800
  - Developed by Carl F. Gauss to fit astronomical observations to Kepler’s Laws of Planetary Motion
  - Difficult method using hand computations
    - Easily performed with a computer
  - Provides the most probable solution for a set of observations!

How Many Solutions?

- Arbitrary methods of adjustment, like the compass rule, often provide different solutions.
  - Dependent on method of adjustment
  - Often redundant observations not included in adjustment.
- Least squares has only one solution
  - Most probable solution
  - All the observational data are used



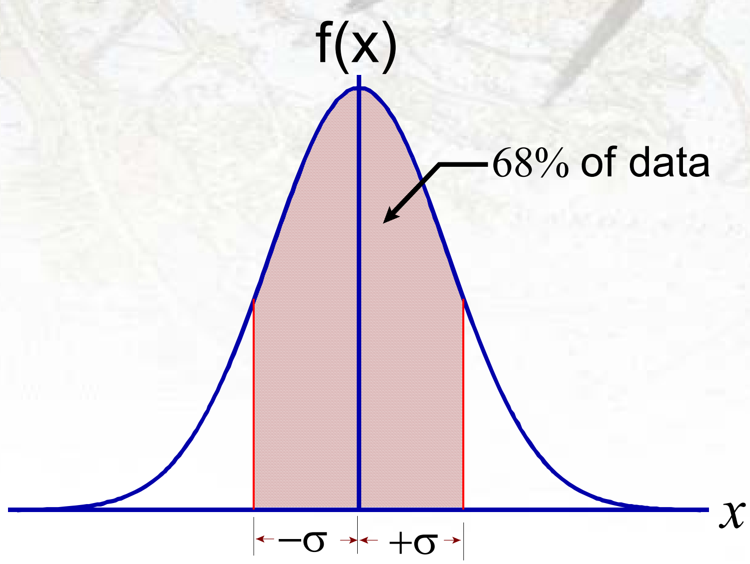
Development of Least Squares

- Suppose there are "n" observations $z_1, z_2, \ldots, z_n$
- Let M be the *most probable value* for z
- The residuals are:

$$M - z_1 = v_1$$
$$M - z_2 = v_2$$
$$\vdots$$
$$M - z_n = v_n$$

Normal Distribution Curve

- Normal distribution curve defined by $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$



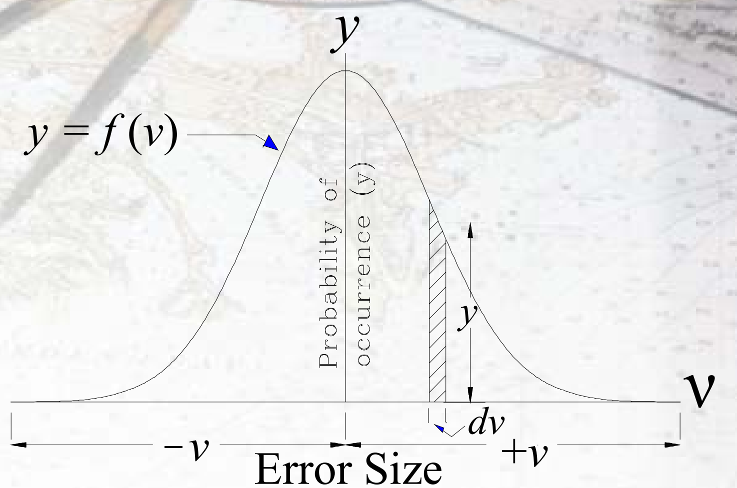
Reworking the Normal Density Function

- Normal density function is $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$
- Substituting $h = \frac{1}{\sigma\sqrt{2}}$ and $K = \frac{h}{\sqrt{\pi}}$
- and v for x, the normal density function becomes

$$y_i = \frac{1}{\sigma\sqrt{2\pi}} e^{-v_i^2/2\sigma^2} = Ke^{-hv_i^2}$$

To Determine a Probability of a Particular Event

- The probability of any observation can be determined by the area under the normal distribution curve
  - Use an *infinitesimally small* region bounding the error, dv
  - Multiply by the value of *y*
  - Note: This is theoretical; i.e. not practical



Probability of Each Residual

$$P_1 = y_1 \Delta v = K e^{-h^2 v_1^2} \Delta v$$
$$P_2 = y_2 \Delta v = K e^{-h^2 v_2^2} \Delta v$$
$$\vdots$$
$$P_n = y_n \Delta v = K e^{-h^2 v_n^2} \Delta v$$

So what is the probability of P_1, to P_n occurring simultaneously?

Probability

- **Probability of the simultaneous occurrence of all the events is the <u>product of all the individual probabilities</u>.**

- So

$$P = \underbrace{(Ke^{-h^2 v_1^2}\Delta v)(Ke^{-h^2 v_2^2}\Delta v)\cdots(Ke^{-h^2 v_n^2}\Delta v)}_{n}$$

$$P = K^n (\Delta v)^n e^{-h^2(v_1^2 + v_2^2 + \cdots + v_n^2)}$$

Development

$$P = K^n(\Delta v)^n e^{-h^2(v_1^2+v_2^2+\cdots+v_n^2)}$$

- Recall $h = \frac{1}{\sigma\sqrt{2}}$ and $K = \frac{h}{\sqrt{\pi}}$
- K,n, Δv, e, and h are all constants
  - Only thing we can change is sum of the squared residuals
- To maximize P, we must minimize the power of e;
- i.e. Minimize the sum of the squared residuals!

Fundamental Principle of Least Squares

- "The most probable value for a measured quantity is the value that renders the sum of the squared residuals a minimum."

Back to Problem

- Thus, we must minimize Σv^2

$$\sum v^2 = (M - z_1)^2 + (M - z_2)^2 + \cdots + (M - z_n)^2$$

- To minimize, take the first derivative and set the resulting equation equal to 0.

$$\frac{\partial \sum v^2}{\partial M} = 2(M - z_1)(1) + 2(M - z_2)(1) + \cdots + 2(M - z_n)(1)$$
$$= nM - (z_1 + z_2 + \cdots + z_n)$$
$$= 0$$

- So, $M = (z_1 + z_2 + \cdots + z_n)/n$, the mean of a simple set of data!

What is Least Squares?

- Consist of two components
  - Functional model
    - Set of mathematical equations that define the observations
  - Stochastic model
    - Weighting model of observations
    - An advantage of lsq over other adjustment methods is that it let’s the user provide differences in qualities of observations
  - Both models must be correct to get the most probable value for the solution

What is Least Squares?

- Just as in the compass rule adjustment, there are more equations than there are unknowns, which provides our checks
- The only difference between what you did in your algebra class and in least squares is how the equations are formed
  - And the fact that the number of equations and unknowns is typically large, and thus solved with computers using matrices, or an icon on your computer screen.

Functional Model

- Optical observations
  - Leveling: $E_j - E_i = \Delta Elev_{ij} + v_{\Delta Elev_{ij}}$
  - Traversing
    - Distance: $\sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} = l_{ij} + v_l$
    - Azimuth: $\tan^{-1}\left(\frac{x_j - x_i}{y_j - y_i}\right) + C = \alpha_{ij} + v_\alpha$
    - Angles: $\alpha_{IF} - \alpha_{IB} = \measuredangle BIF + v_{\measuredangle}$

GPS Functional Model

– $\Phi_{L1}^{j}(t) + v_{\Phi_{L1}} = \frac{1}{\lambda_{L1}}\rho^{j}(t) + f_{L1}\Delta\delta^{j} + N_{L1} - f_{L1}\delta^{Iono} + f_{L1}\delta^{Trop}$

– $\Phi_{L2}^{j}(t) + v_{\Phi_{L2}} = \frac{1}{\lambda_{L2}}\rho^{j}(t) + f_{L2}\Delta\delta^{j} + N_{L2} - f_{L2}\delta^{Iono} + f_{L2}\delta^{Trop}$

– $\Phi_{L5}^{j}(t) + v_{\Phi_{L5}} = \frac{1}{\lambda_{L5}}\rho^{j}(t) + f_{L5}\Delta\delta^{j} + N_{L5} - f_{L5}\delta^{Iono} + f_{L5}\delta^{Trop}$

Clock biases | Integer ambiguities | Refraction components

– where

- $\Phi_{L1}^{j}(t)$, $\Phi_{L2}^{j}(t)$, & $\Phi_{L5}^{j}(t)$ are the phase-shift observations from the receiver to satellite *j* at epoch *t* for the *L*1, *L*2, *L*5 frequencies
- $\rho^{j}(t) = \sqrt{(X^{j} - X_{R})^{2} + (Y^{j} - Y_{R})^{2} + (Z^{j} - Z_{R})^{2}}$ where $(X,Y,Z)^{j}$ and $(X,Y,Z)_{R}$ are the geocentric coordinates of satellite *j* and receiver *R* at epoch (*t*)

Integer Ambiguities

- Today, *N* is determined using the LAMBDA method
  - LAMBDA is an acronym for **L**east-squares **AMB**iguity **D**ecorrelation **A**djustment
- So, carrier phase-shift (kinematic, static, network) surveys all “initialize” by using a least squares adjustment

Stochastic Model

- The stochastic model lets the user control the amount of correction that will occur in each observation
  - This is the power that least squares gives you
  - The important item here is that there **must** be some valid statistical reason for determining the uncertainty in the observations and the control

Stochastic Model: Differential Leveling

- Differential leveling:
  - $\sigma_{\Delta h} = D\sqrt{2N(\sigma_{r/d}^2 + \sigma_{\alpha}^2)}$
    - Where
      - *D*, the typical sight length
      - *N*, the number of setups
      - $\sigma_{r/d}$, the ability to read the rod per unit sight length
      - σ_{α}, the ability of the automatic compensator to find horizontal – misleveling error in radians

Example from Spec Data Sheet

- Uncertainty in reading, $\sigma_{r/d}$
- Compensator setting accuracy, σ_α

HEIGHT MEASUREMENT	
Accuracy[1] using standard Invar staff	0.3 mm
Accuracy[2] using standard staff	1.0 mm
DISTANCE MEASUREMENT	
Accuracy[3]	15 mm at 30 m
MEASUREMENT RANGE	
Minimum range	1.8 m
Maximum range[4]	110 m
Measurement time	Typically 2.5 sec
AUTOFOCUS	
Working range	✗
Time to focus	✗
OVERVIEW CAMERA	
Field of view	✗
Frame rate	✗
Focus	✗
DIGITAL COMPASS	
Accuracy[6]	✗
COMPENSATOR	
Working range	± 9'
Accuracy[6]	0.3''
Magnetic field sensitivity[5]	≤ 1''

Stochastic Model: EDM Observations

- Uncertainties are
  - Instrument constant error, a; E.g. 5 mm, 3 mm, 2 mm, etc.
  - Instrument scalar error, b; E.g. 5 ppm, 3 mm, 2 ppm, etc.
  - Instrument miscentering error, σ_i
  - Reflector miscentering error, σ_r
- Error in observed distance is

$$\sigma_D = \sqrt{\sigma_i^2 + \sigma_r^2 + a^2 + (D \times b \text{ ppm})^2}$$

Example

- Distance of 453.87 ft is observed using an EDM with a specified accuracy of 3 mm (~0.01 ft) + 3 ppm. The instrument set up error is estimated to be ±0.003 ft and the reflector set up error is estimated to be ±0.01ft.
- What is the estimated uncertainty in the distance?

$$\sigma_D = \sqrt{0.003^2 + 0.01^2 + 0.01^2 + \left(\frac{3}{1{,}000{,}000}453.87\right)^2} = \pm 0.015 \text{ ft}$$

Stochastic Model: Angles

- Sources of Errors
  - Pointing and reading of the instrument, σ_{ISO}
    - σ_{ISO} based on manufacturer's specs for instrument
  - Miscentering of the instrument, σ_i
  - Miscentering of the target, σ_t
  - Leveling of the instrument, σ_α
    - Only significant when vertical angles are large

Stochastic Model: Pointing and Reading

- Pointing and reading error: $\sigma_{\alpha_{pr}} = \frac{2\sigma_{ISO}}{\sqrt{n}}$
  - Where σ_{ISO} is the instruments ISO specification
  - n, number of repetitions of an angle

σ_{ISO} from tech. specs.

Angle Measurement	
	Min. Resolution/Accuracy
ES-101	0.5″/1″
ES-102	1″/2″
ES-103	1″/3″
ES-105	1″/5″
ES-107	1″/7″

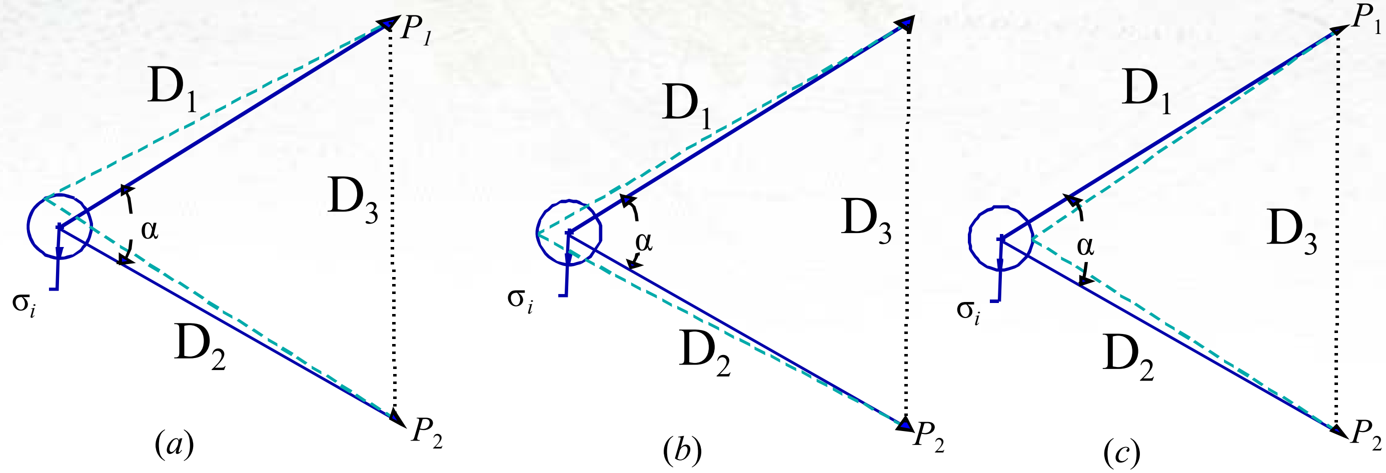
Stochastic Model: Pointing and Reading

- An angle is observed 2 times with a total station having a manufacturer's ISO specification of ±3″
- What is the error in the angle due to pointing and reading of the angle?

$$\sigma_{\alpha_{pr}} = \frac{2(3'')}{\sqrt{2}} = \pm 4.2''$$

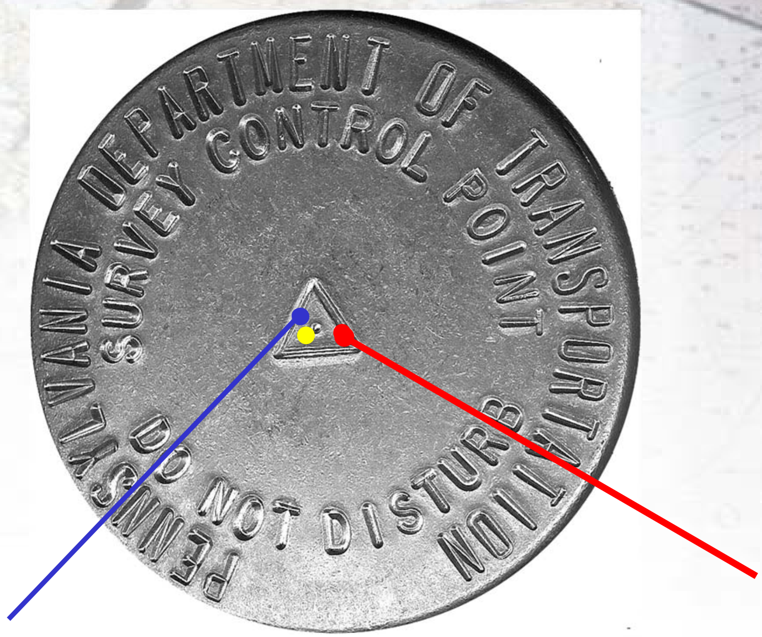
Stochastic Model: Instrument Miscentering

- Error in every direction observed
- Errors <u>larger</u> for <u>shorter</u> sight distances
- Error will only be "seen" in a second setup of instrument, i.e. a resurvey



How Setup Errors Affect Coordinates

- Instrument setup
- Foresight direction and distance
- Backsight direction and distance
- End result is multiple locations for station
  - Even though it is a systematic error, it will appear to be a random error in the adjustment



Sighting Errors

- Never sight your retro-reflector for an angle
  - Causes line of sight to shift if reflector not perpendicular to line of sight exactly
  - Sight an angel face or a chaining pin, supported plumb bob string, etc.

Stochastic Model: Instrument Miscentering

- Estimated error in an angle is

$$\sigma''_\alpha = \frac{D_3}{D_1 D_2 \sqrt{2}} \sigma_i \times 206264.8\,''/\text{radian}$$

  - Where
    - D_3 is the distance between the backsight (D_1) and foresight (D_2) targets

Stochastic Model: Instrument Miscentering

- Observed angle is 50°00'00"
- Estimated instrument miscentering error, σ_i, is 0.005 ft.
- Backsight distance (D_1) is 250 ft
- Foresight distance (D_2) is 450 ft
- What is the estimated error in the angle due to instrument miscentering?

Stochastic Model: Instrument Miscentering

- Using cosine law

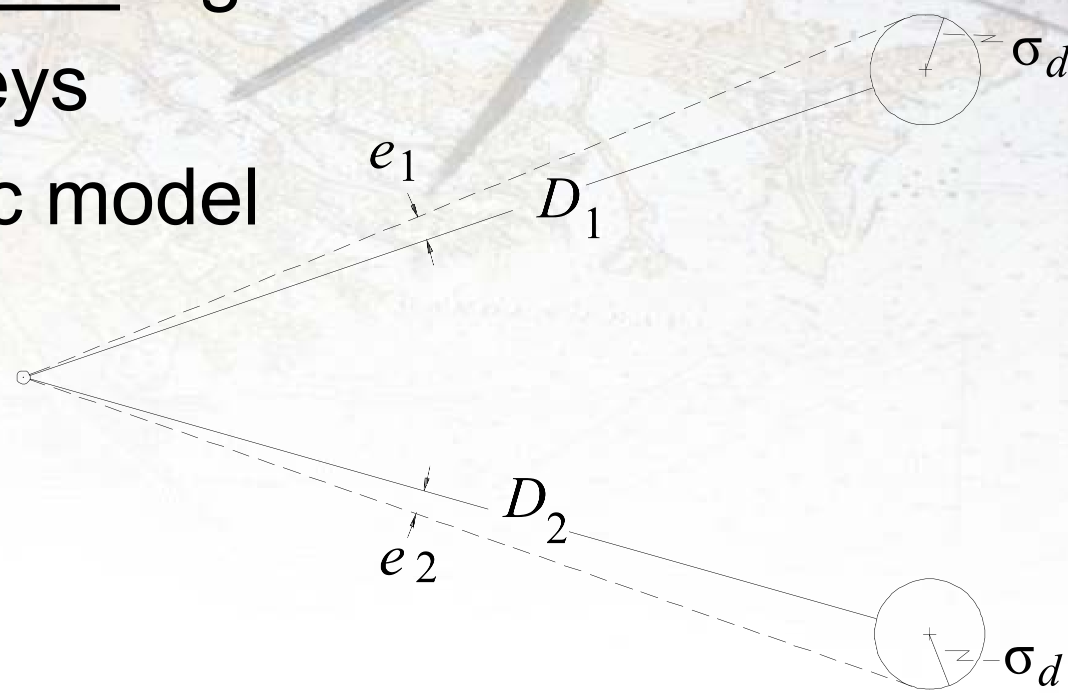
$$D_3 = \sqrt{250^2 + 450^2 - 2(250)(450)\cos 50^\circ}$$
$$= 346.95 \text{ ft}$$

- So

$$\sigma''_\alpha = \frac{346.95}{250(450)\sqrt{2}}(0.005)(206{,}264.8) = \pm 2.2''$$

Stochastic Model: Target Miscentering

- <u>Larger</u> errors on <u>shorter</u> sight distances
- Only seen in resurveys
- But part of stochastic model



Stochastic Model: Target Miscentering

- Target centering error:

$$\sigma''_{\alpha_t} = \pm \frac{\sqrt{D_1^2 + D_2^2}}{D_1 D_2} \sigma_t (206{,}264.8\,''/\text{radian})$$

  - Where D_1 and D_2 are the sight lengths to target 1 and 2, σ_t the estimated error in target miscentering

Stochastic Model: Target Miscentering

- σ_t is estimated as ±0.01 ft
- Backsight distance is 250 ft
- Foresight distance is 450 ft
- What is the estimated error in the angle due to target miscentering?

$$\sigma_t = \frac{\sqrt{250^2 + 450^2}}{250(450)} 0.01\,(206264.8''^{/\mathrm{rad}}) = \pm 9.4''$$

Stochastic Model: Angles

- Overall uncertainty in this one angle is

$$\sigma''_{\alpha} = \sqrt{4.2''^2 + 2.2''^2 + 9.4''^2} = \pm 10.5''$$

An Aside

- Assume that a 7-sided traverse was observed with these uncertainties. What would be a reasonable maximum allowable misclosure?
- First: Instrument and target miscentering are only discovered on a resurvey, so only the pointing and reading error is applicable.
- So at 95%, the maximum allowable error is

$$\sigma_{\sum \measuredangle} = 1.96\sqrt{4.2^2 + 4.2^2 + 4.2^2 + 4.2^2 + 4.2^2 + 4.2^2 + 4.2^2} = \pm 22''$$

How to Handle Control

- What do you do when there is more than the minimum control?
- Minimum control is:
  - Plane survey:
    - 1 control point; 1 course fixed in direction
  - Differential level:
    - 1 benchmark
  - GNSS survey:
    - 1 horizontal and 1 vertical control point

How to Handle Control

- Procedure
  - Adjust observations using minimum control only
    - Remove any blunders from data then
    - Note reference variance and redundancies
  - Adjust observations using all control
    - Note reference variance and redundancies
  - Do an *F* test to determine if the two "sample" reference variances, S_0^2, are statistically the same

How to Handle Control

- If H_0 is rejected, then either there is a
  - Blunder in the control
    - For a control survey, you should have a minimum of 3 control points. Why?
  - Systematic error in the observations
    - For example
      - When using SPCS, failure to reduce observations to the adjustment surface
      - Constant error in distance (e.g. EDM/reflector constant)
  - Both!

Control Observations?

- Can you adjust your control
  - With a compass rule adjustment? No, must be fixed
  - With a least squares adjustment? Yes
    - Control can be weighted as appropriate
      - Example: Suppose control azimuth is determined by GNSS survey where stated accuracy is listed as ±21″
        - » Weight azimuth using ±21″, and it may move to fit observations
      - Control station coordinates can also be weighted as appropriate

Adjusting Control

- Control is determined from a previous survey, so it also contains errors
  - Key is to determine appropriate standard deviations to weight control
  - NGS list horizontal and vertical control uncertainties

Functional Model: Control

- A simple set of equations

$$X' = X + v_X$$
$$Y' = Y + v_Y$$
$$Z' = Z + v_Z$$

  – where
    - X', Y', Z' are the adjusted control from your adjustment
    - X, Y, Z are the reported control from a previous survey or the NGS
    - v_X, v_Y, v_Z are the residuals after the adjustment

Stochastic Model: Control

- If control is from a previous adjustment, enter the standard deviations for the coordinates (68%)
- If it is still listed by precision then compute value based on precision

Stochastic Model: Control

- For control that is part of the NSRS active control and their reference marks
  - Standard deviations in
    - Northing or *y*, SD_N: ±7.8 mm
    - Easting or *x*, SD_E: ±6.5 mm
    - Height, SD_h: ±2.95 cm
      - Note: ellipsoid height + Geoid height error
  - Will discuss near end of workshop

```
        National Geodetic Survey,   Retrieval Date = SEPTEMBER 14, 2017
DF5874 ***********************************************************************
DF5874  CORS          -  This is a GPS Continuously Operating Reference Station.
DF5874  DESIGNATION -  HARRISBURG CORS ARP
DF5874  CORS_ID      -  GTS1
DF5874  PID          -  DF5874
DF5874  STATE/COUNTY-  PA/DAUPHIN
DF5874  COUNTRY      -  US
DF5874  USGS QUAD    -  HARRISBURG EAST (1987)
DF5874
DF5874                        *CURRENT SURVEY CONTROL
DF5874  ______________________________________________________________________
DF5874* NAD 83(2011) POSITION- 40 15 07.78723(N) 076 49 34.83797(W)   ADJUSTED
DF5874* NAD 83(2011) ELLIP HT-    89.299 (meters)        (08/??/11)   ADJUSTED
DF5874* NAD 83(2011) EPOCH    -  2010.00
DF5874  ______________________________________________________________________
DF5874  GEOID HEIGHT     -          -34.327 (meters)                  GEOID12B
DF5874  NAD 83(2011) X  -  1,110,966.806 (meters)                  COMP
DF5874  NAD 83(2011) Y  - -4,746,446.484 (meters)                  COMP
DF5874  NAD 83(2011) Z  -  4,099,452.422 (meters)                  COMP
DF5874
DF5874  Network accuracy estimates per FGDC Geospatial Positioning Accuracy
DF5874  Standards:
DF5874         FGDC (95% conf, cm)     Standard deviation (cm)     CorrNE
DF5874            Horiz  Ellip          SD_N   SD_E   SD_h      (unitless)
DF5874  -------------------------------------------------------------------
DF5874  NETWORK     1.76    5.78          0.78   0.65   2.95      0.00578500
DF5874  -------------------------------------------------------------------
```

Stochastic Model: Passive Control

- Most passive control currently listed by precision such as (1:20,000)
- Compute uncertainty based on nearest control

$$\sigma_P = \sigma_c\sqrt{2}$$

  – where
    - σ_P is the uncertainty based on the reported precision
    - σ_c is the estimated uncertainty for the *x* and *y* coordinates

$$\sigma_c = {}^{\sigma_P}\!/_{\sqrt{2}}$$

Stochastic Model: Passive Control

- Assume an adjustment has two control stations which are 8426.78 ft apart. One is listed as 1:10,000 precision and the second is listed as 1:50,000. What are reasonable estimates for the uncertainties of the control coordinates?
  - $\sigma_{P1} = 1\!:10{,}000(8426.78) = \pm 0.84\text{ ft}$
  - $\sigma_{P2} = 1\!:50{,}000(8426.78) = \pm 0.17\text{ ft}$
  - $\sigma_{c1} = \frac{0.84}{\sqrt{2}} = \pm 0.59\text{ ft}$
  - $\sigma_{c2} = \frac{0.17}{\sqrt{2}} = \pm 0.12\text{ ft}$
    - enter these values for the *x* and *y* coordinates in the adjustment

Stochastic Model: GPS

- First: You don’t have control over much!
- Second: When do interact with the adjustment, such as in a static survey, the stochastic model is set by the prior lsq adjustments you had no control over
- Third: The manufacturer’s all do this incorrectly!
- Finally: Until statistical models are developed to estimate error in sat. positioning, refraction, etc. GPS will never achieve it best results or yield true MPVs

Stochastic Model: GNSS Setup Errors

- Setup errors can be large!
  - Follow a Rayleigh's distribution
  - Assume a 8′ bubble that is within 1 division of being centered
    - 8′ per 2-mm division
  - Horizontal error: $E_r = \sqrt{(f_d \mu \sigma_r)^2 + (R\sigma_\theta)^2}$
    - where
      - f_d is the fractional division of miscentering, 1 division
      - μ is the sensitivity of the bubble, 8′
      - σ_r is the estimated error in the length of the rod, 3 mm
      - R is the height of the receiver
      - $\sigma_\theta = f_d \mu \sqrt{0.5(4 - \pi)}$

Stochastic Model: GNSS Setup Errors

Assuming a rod length of 2 m

- Fixed-height tripod, which has a $8'$ circular level vial, $8' \simeq 0.002327$ radians
- $\sigma_\theta = 1(0.002327)\sqrt{0.5(4-\pi)} = \pm 0.001525$ radians
- $E_r = \sqrt{(f_d \mu \sigma_r)^2 + (R\sigma_\theta)^2}$

$$= \sqrt{(1(0.002327)\ 3\ \text{mm})^2 + \big(2000(0.001525)\big)^2}$$

$$= \sqrt{0.006981^2 + 3.05^2} = \pm 3.05\ \text{mm} = \pm 0.01\ \text{ft}$$

Stochastic Model: GNSS Setup Errors

- If you use a typical mapping-grade rod with a 40′ bubble
  - Assume the other values the same
  - 40′ in radians is 0.0116355 radians
- $\sigma_\theta = 1(0.0116355)\sqrt{0.5(4-\pi)} = \pm 0.007623 \text{ radians}$
- $E_r = \sqrt{(f_d \mu \sigma_r)^2 + (R\sigma_\theta)^2}$

$$= \sqrt{(1(0.0116355)\ 3\text{ mm})^2 + \left(2000(0.007623)\right)^2}$$

$$= \sqrt{0.035^2 + 15.25^2} = \pm 15.2\text{ mm} = \pm 0.05\text{ ft}$$

- CHECK YOUR ROD BUBBLE SPECIFICATIONS!

Stochastic Model

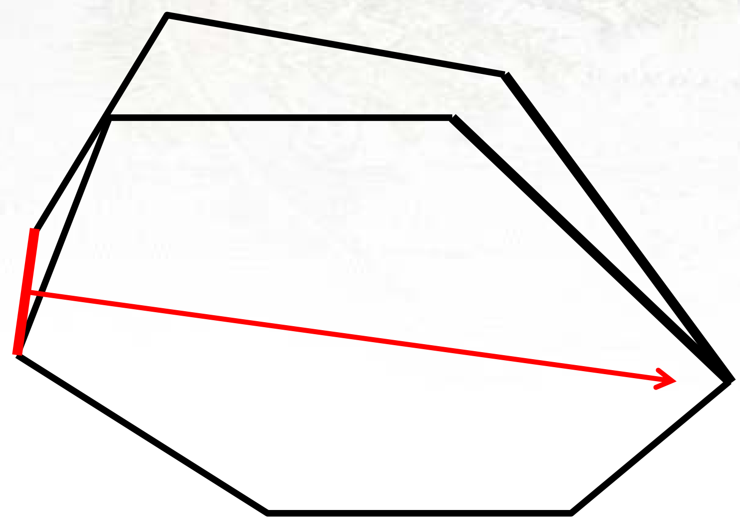
- At least when doing a combined optical-satellite survey consider coming up with reasonable values for uncertainties in your angles and distances; i.e. a reasonable average value.
  - Note: you can usually only enter one value for each observation type
  - Enter reasonable values for setup errors and control uncertainties

So…

- You have run your least squares adjustment or obtained your positions using GNSS, what is next???

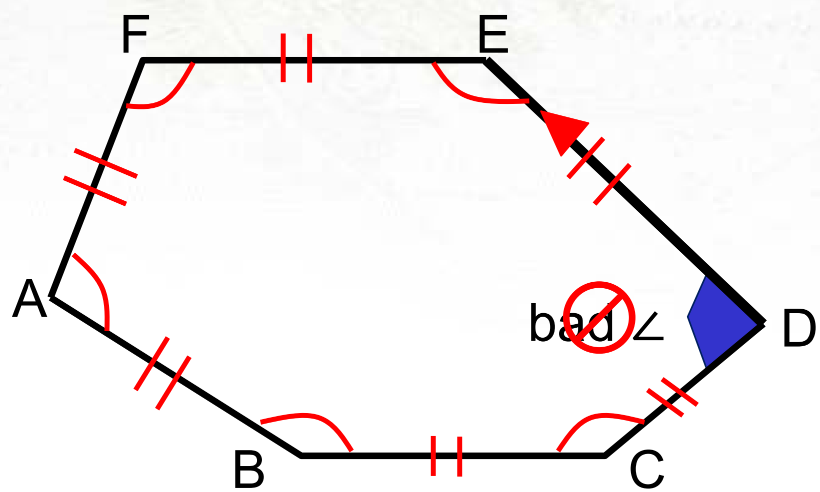
Before the Adjustment: Graphical Methods

- Angular blunders rotate the remaining courses of a traverse.
- The perpendicular bisector of the misclosure line (chord on a circle) will point close to the angle with a blunder.



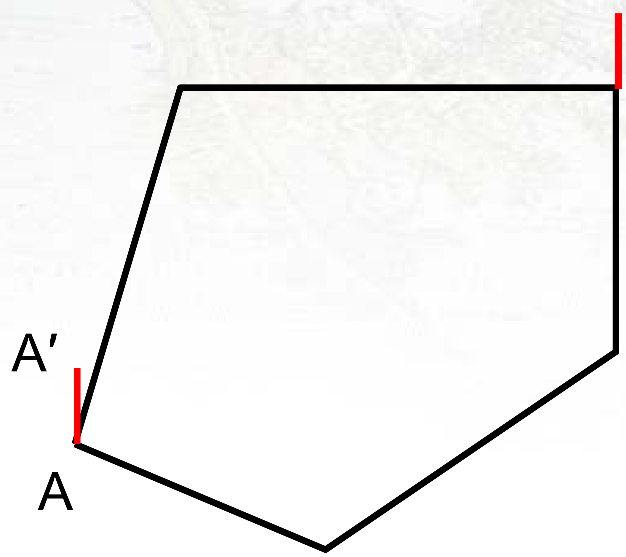
Before the Adjustment: Graphical Methods

- Start your computations at the station with the suspected “bad” angle to eliminate the angle from the computations
  - Assume starting azimuth *DE* and use *DE* length to compute coordinates @ *E*
  - Use ∠*E* and distance *EF* to compute coordinates for *F*
  - Use ∠*F* and distance *FA* to compute coordinates for *A*
  - Use ∠*A* and distance *AB* to compute coordinates for *B*
  - and so on



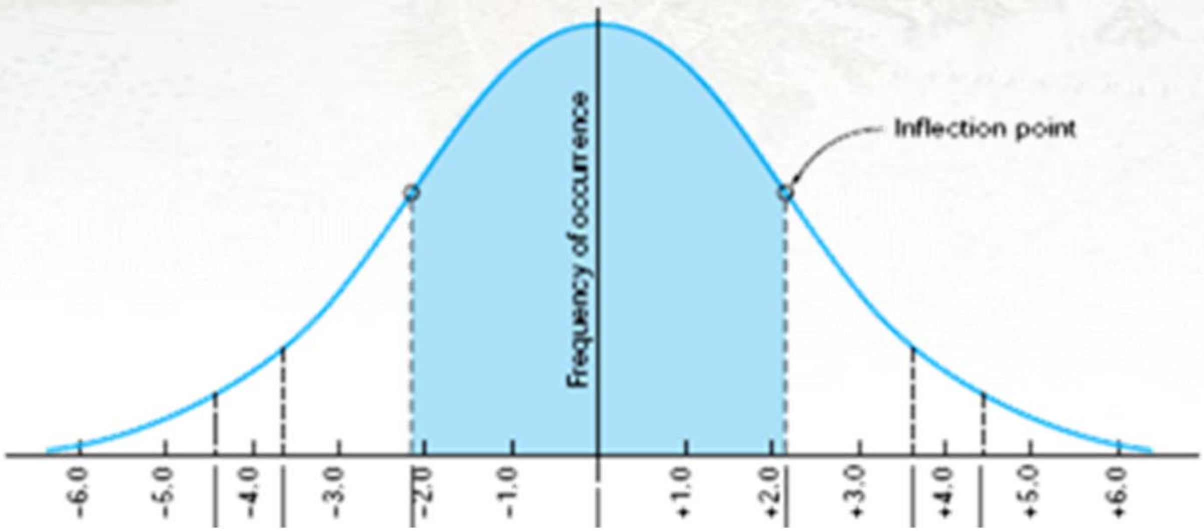
Graphical Methods

- Distance blunders translate the remaining courses of a traverse
  - Bearing of misclosure line will be close to bearing of line with blunder
  - Must check field notes or reobserve



Recall: Concepts of Normally Distributed Data

- Our data should be normally distributed
- What does the plot tell us?
  – Small errors occur more frequently than large errors
    - Large errors seldom occur

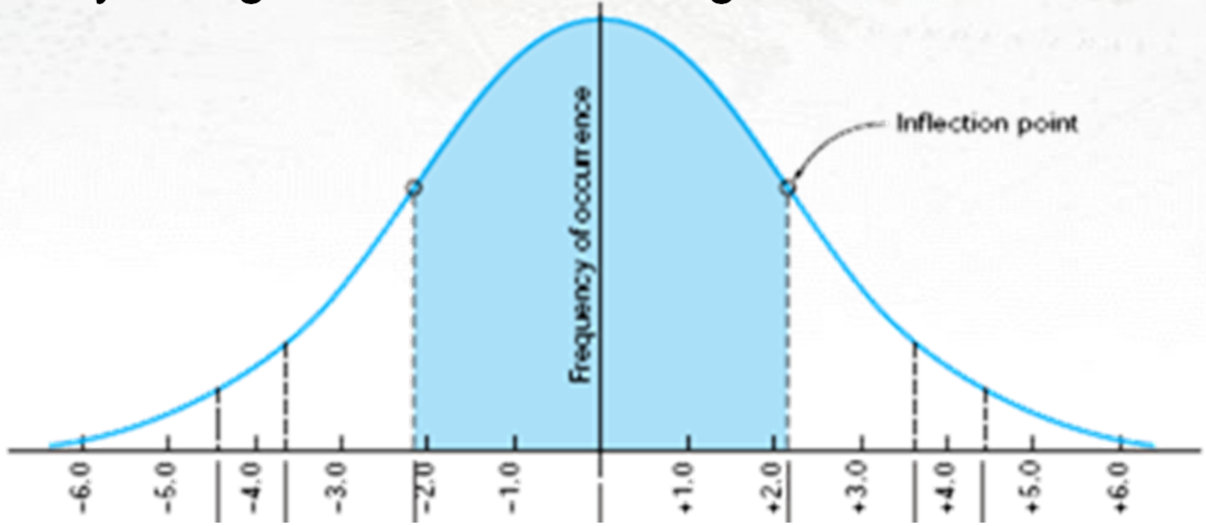


Concepts of Normally Distributed Data

- As a surveyor we can choose to remove all observations outside the 95% or 99% range – your call
- We call these blunders knowing that 5% to 1% of the time they will actually be legitimate observations with large random errors, otherwise known as outliers
  - This will probably remove most blunders

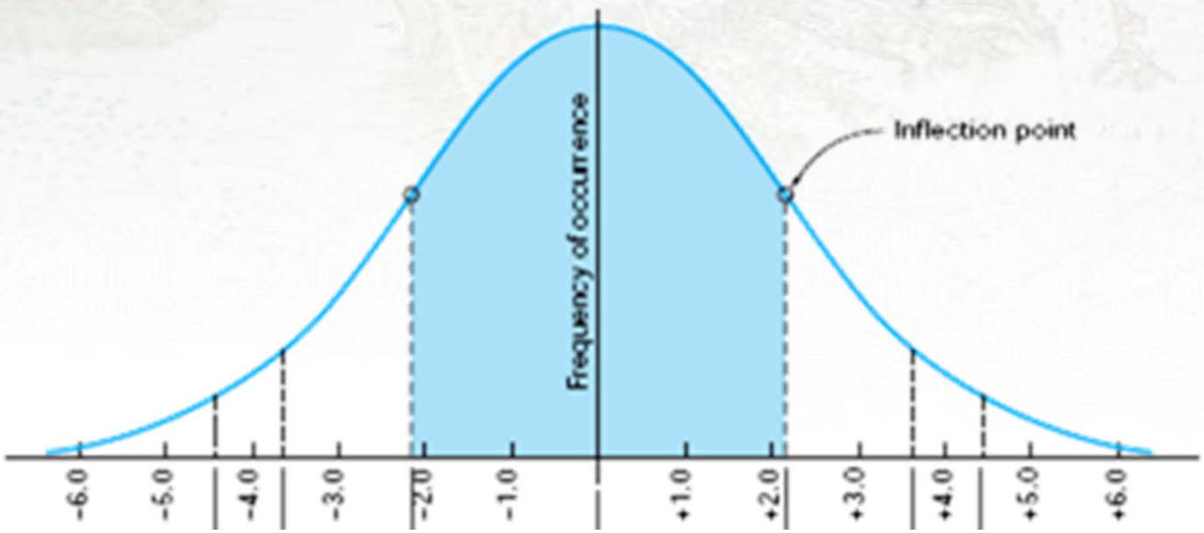
Recall: Concepts of Normally Distributed Data

- Tend to be equally distributed
  - There should be as many positive residuals as there are negative.
    - If residuals are heavily lopsided (positive versus negative) then typically a large blunder is causing this



Concepts of Normally Distributed Data

- The signs of errors tend to occur in series. That is, ++−−−++++−−+−−−, and so on.



Useful Probabilities

Symbol	Multiplier	Percentage
E_{50}	0.675 σ	50
E_{90}	1.645 σ	90
E_{95}	1.960 σ	95
E_{99}	2.576 σ	99
$E_{99.7}$	2.965 σ	99.7

- Note: For a normally distributed population, 99.7% of data should be within $\pm 2.965\sigma$ ($\sim 3\sigma$) of the mean
- Use these values to define large error

Example 16.2

- Using ADJUST analyze the data for Example 16.2

Example 16-2.adat

Analyze Adjusted Distances

Station Occupied	Station Sighted	Distance	v	S
Q	R	1,639.978	-0.0384	0.0236
R	S	1,320.019	0.0176	0.0228
S	T	1,579.138	0.0155	0.0235
T	Q	1,664.528	0.0039	0.0250
Q	S	2,105.953	-0.0087	0.0232
R	T	2,266.045	0.0104	0.0241

- Residuals not balanced but only off by one. Not bad.
  - Remember, this is a sample of data & only part of data set
- Largest $|v|$ is 0.0384. Its *a priori* $S = \pm 0.026$
  - At 99%, Multiplier is about 2
  - So, acceptable range is 3 (0.026) = ±0.078 ft.
  - $|v_{QR}| < 0.078$

Stats

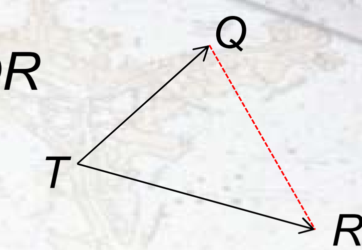
Analyze Adjusted Angles

BS	Occ	FS	Angle	V	S"
R	Q	S	38°48'52.8"	2.08"	2.59
S	Q	T	47°46'13.1"	0.70"	2.89
T	Q	R	273°24'54.1"	-2.37"	3.56
Q	R	S	269°57'34.0"	0.64"	3.36
R	S	T	257°32'57.3"	0.50"	3.71
S	T	Q	279°04'34.5"	3.32"	3.45
S	R	T	42°52'52.3"	1.34"	2.70
S	R	Q	90°02'26.0"	-0.74"	3.36
Q	S	R	51°08'41.3"	-3.73"	2.94
T	S	Q	51°18'21.4"	5.23"	3.02
Q	T	R	46°15'20.5"	18.52"	2.70
R	T	S	34°40'05.0"	-0.74"	2.55

Note: 4 “−” residuals
8 “+” residuals

Analyze Adjusted Angles

- Note size of residual for angle *QTR*
  - Note that an error in $\angle QTR$ will directly affect distance *QR*
- a priori value for $S_{QTR} = \pm 4.0''$
  - 99% acceptable range is $3(4.0) = \pm 12.0''$
  - $18.5'' = v_{QTR} > 12.0''$, so $\angle QTR$ might be a blunder!
- Remove it from data set, run, and analyze adjustment again



Second Adjustment Attempt

Adjusted Distances

Occupied	Sighted	Distance	V	S
Q	R	1,640.008	-0.0081	0.0060
R	S	1,320.006	0.0054	0.0055
S	T	1,579.133	0.0099	0.0056
T	Q	1,664.514	-0.0097	0.0060
Q	S	2,105.966	0.0039	0.0056
R	T	2,266.034	-0.0014	0.0058

Note: 3 −residuals and 3 +residuals

Second Adjustment Attempt

Adjusted Angles

Backsighted	Occupied	Foresighted	Angle	V	S"
R	Q	S	38°48'50.2"	-0.45"	0.64
S	Q	T	47°46'11.7"	-0.73"	0.69
T	Q	R	273°24'58.1"	1.58"	0.89
Q	R	S	269°57'34.7"	1.31"	0.80
R	S	T	257°32'56.9"	0.11"	0.88
S	T	Q	279°04'30.3"	-0.91"	0.87
S	R	T	42°52'52.6"	1.58"	0.64
S	R	Q	90°02'25.3"	-1.41"	0.80
Q	S	R	51°08'44.5"	-0.53"	0.73
T	S	Q	51°18'18.6"	2.43"	0.74
R	T	S	34°40'04.3"	-1.37"	0.61

Note: 6 "−" residuals and 5 "+" residuals
Total: 9 "−" residuals and 8 "+" residuals

Second Adjustment Attempt

Post-Adjustment Statistics

```
****************************************
         Adjustment Statistics
****************************************
                 Iterations = 2
              Redundancies = 12

       Reference Variance = 0.1243
              Reference So = ±0.35

Failed to pass χ² test at 95.0% significance level!
              χ² lower value = 4.40
              χ² upper value = 23.34
                 Convergence!
```

χ^2 Test

- Reference variance computed as

$$S_0^2 = \frac{\sum wv^2}{redundancies} = \frac{V^{\mathrm{T}}WV}{redundancies}$$

- **Failed because reference variance was too small**! Should be close to 1!
  - Causes: Residuals too small!
    - Stochastic model incorrect
      - » That is, the observations are more accurate than the a priori estimates in the stochastic model
  - Solution
    - Accept adjustment as is
    - Adjust the stochastic model

Post-Adjustment Statistics: Tau Criterion

- Used in automatic blunder detection
- Can catch small blunders in data
- Is sensitive to the stochastic model

Statistical Methods for Isolating Blunders

- Create the covariance matrix of the residuals
- Compute standard deviations for the residuals
  - If we know S_v, then we can predict range of acceptable *v's*
- Use *tau* criterion (τ test) to statistically isolate blunders

Tau Statistic Theory

- Standardized residual is computed as

$$\bar{v}_i = \frac{v_i}{\sqrt{q_{ii}}}$$

  - where q_{ii} is the diagonal of the i'th diagonal of the cofactor matrix for the residuals, Q_{vv}

- A computed parameter divided by its standard deviation is a τ statistic

$$\tau_i = \frac{v_i}{S_0\sqrt{q_{ii}}} = \frac{v_i}{S_{v_i}}$$

Tau Statistic

- Baarda developed method using *t* statistic (data snooping) but Alan J. Pope (from NGS) suggested the *tau statistic* be used where τ is computed as

$$\tau_i = \frac{v_i}{S_0\sqrt{q_{v_i v_i}}}$$

- With the rejection criteria being $|\tau_i| > \tau_{\frac{\alpha_0}{2}}$
- The test are generally performed at α = 0.001, or 99.9% probability.

Tau Statistic

- Works very well for small blunders
- Standardized residual is compared against a statistical quantity
  - Observation may be a blunder if standardized residual > rejection level
- Software either automatically removes observations with blunders or highlights them

Manual Steps

1. Locate all measurements that qualify for rejection.
2. Reject the single observation with the largest standardized residual.
3. Repeat adjustment and steps 1 & 2 until all observations the qualify for rejection are rejected.
4. Reinsert observations that were rejected into adjustment in a *one-at-a-time* fashion to see if they are still rejected.

Example 16.2

- Using ADJUST analyze the data for Example 16.2

Example 16-2.adat

Using Tau Criterion for Blunder Rejection

Station Backsighted	Station Occupied	Station Foresighted	Angle	V	Std.Res.	Red.#
R	Q	S	38°48'52.8"	2.08"	0.58	0.810
S	Q	T	47°46'13.1"	0.70"	0.20	0.762
T	Q	R	273°24'54.1"	-2.37"	-0.64	0.702
Q	R	S	269°57'34.0"	0.64"	0.16	0.768
R	S	T	257°32'57.3"	0.50"	0.13	0.717
S	T	Q	279°04'34.5"	3.32"	0.86	0.733
S	R	T	42°52'52.3"	1.34"	0.34	0.821
S	R	Q	90°02'26.0"	-0.74"	-0.19	0.747
Q	S	R	51°08'41.3"	-3.73"	-0.98	0.787
T	S	Q	51°18'21.4"	5.23"	1.52	0.740
Q	**T**	**R**	**46°15'20.5"**	**18.52"**	**5.20**	**0.792***
R	T	S	34°40'05.0"	-0.74"	-0.20	0.815

Tau criterion used.

Critical tau value: 3.286 at 0.0010 level of significance.

Possible blunder in observations with **Std.Res. > 4.869**

Another Example with Automatic Blunder Removal

Subnetwork JASH, WIL1 B, SHUPPSKI, ... (Horizontal Minimal Constraint + Vertical Minimal Constraint)

Type	Adjusted Points	Fixed Points	Weighted Points	Equations (Used/Rejected)	UWE	UWE Bounds
				GPS		
Horz	6	1	0	17	3.04	[0.64,1.38]
Vert	6	1	0	17	2.11	[0.51,1.54]

- Software displays how many equation used versus the number rejected
- Results of χ^2 test

Another Example of Automatic Blunder Detection

Subnetwork JASH, WIL1 B, SHUPPSKI, ... (Horizontal Minimal Constraint + Vertical Minimal Constraint)

Type	Adjusted Points	Fixed Points	Weighted Points	Equations (Used/Rejected) GPS	UWE	UWE Bounds
Horz	6	1	0	17 / 5	1.23	[0.54,1.50]
Vert	6	1	0	17 / 2	1.47	[0.46,1.59]

Rejected Observations

Name	Type	Residual N(m)	Residual E(m)	Residual H(m)
SHUPPSKI-WIL1 B	GPS	*0.013*	*-0.003*	0.002
JASH-WIL1 B	GPS	*0.010*	*-0.008*	0.001
JASH-WIL1	GPS	*0.009*	*-0.009*	0.002
JASH-WIL1	GPS	*0.003*	*-0.010*	-0.012
SHUPPSKI-WIL1	GPS	*0.009*	*-0.002*	-0.002
JAGER-WIL1	GPS	-0.001	-0.001	*-0.019*
JAGER-WIL1	GPS	-0.001	0.000	*-0.012*

However, auto blunder detection sensitive in incorrect weighting model

How to Check Size of Residuals

Rejected Observations

Name	Type	Residual N(m)	Residual E(m)	Residual H(m)
SHUPPSKI-WIL1 B	GPS	0.013	-0.003	0.002
JASH-WIL1 B	GPS	0.010	-0.008	0.001
JASH-WIL1	GPS	0.009	-0.009	0.002
JASH-WIL1	GPS	0.003	-0.010	-0.012
SHUPPSKI-WIL1	GPS	0.009	-0.002	-0.002
JAGER-WIL1	GPS	-0.001	-0.001	-0.019
JAGER-WIL1	GPS	-0.001	0.000	-0.012

- Use horizontal distance to compute uncertainty estimated by manufacturer

$$\sigma_H = \sqrt{\sigma_i^2 + \sigma_t^2 + \left(H * \frac{ppm}{1{,}000{,}000}\right)^2}$$

- Compute circular error for length using residuals, $S = \sqrt{v_e^2 + v_n^2}$
- Then is $S > 3\sigma_H$

Check Occupations

- Incorrect antenna selection
  - Sometimes true on CORS sites
  - Field crews sometimes not sure of their exact antenna type
  - Sometimes obsolete antenna calibrations are maintained in your manufacturer's file!


```
- <Antenna>
    <Name>TPSGR_3</Name>
    <ID>TPSGR-3</ID>
    <Comment>Obsolete callibrations. Do not use them</Comment>
    <Manufacturer>Topcon</Manufacturer>
    <CallibrationSet>Default absolute</CallibrationSet>
    <Alias>TPSGR3 NONE</Alias>
    <Radius>95.8</Radius>
    <A1>210</A1>
```


I...	Point Name	Antenna ...
	JASH	TPSGR-3
	SHUPPSKI	TPSGR-3
	SHUPPSKI	TPSGR-3
	SHUPPSKI	TPSGR-3
	SHUPPSKI	TPSGR-3
	JASH	TPSGR-3
	JASH	TPSGR-3
	WIL1 B	TPSGR-3
	HAYFIELD SW	GR-3
	WIL1 B	GR-3
	WIL1 B	GR-3

Example of Field Errors

Points | **GPS Occupations** | GPS Obs

I...	Point Na...	Antenna Type	Antenna Heig...	Ant Height Me...
	HAYFIELD SW	TPSGR-3	0.610	Vertical
	HAYFIELD SW	TPSGR-3	0.610	Vertical
	JAGER	TPSGR-3	0.610	Vertical
	JAGER	TPSGR-3	0.610	Vertical
	JASH	TPSGR-3	0.610	Vertical
	JASH	TPSGR-3	0.610	Vertical
	JASH	TPSGR-3	0.610	Vertical

- Note wrong antenna. Should be GR-3.
- Height entered in field as 2 m but software in feet

Different Software but Same Story

GPS Observations

Number of Observations : 11
Number of Outliers : 1

Observation Adjustment (Critical Tau = 2.90). Any outliers are in red.

Obs. ID	From Pt.	To Pt.		Observation	A-posteriori Error (1.96σ)	Residual	Stand. Residual
B5	N 245	KTOM	Az.	350°00'57.1444"	0°00'01.0765"	-0°00'00.1908"	-0.29
			ΔHt.	4.046m	0.009m	-0.021m	-3.15
			Dist.	1076.373m	0.006m	0.005m	1.09
B7	COWBOY	KTOM	Az.	89°24'13.7202"	0°00'00.7512"	-0°00'00.1313"	-0.20

- Software highlights observations that failed test

So Why Process Your Own Data?

- Professionals should know the product they deliver to a client
- Using OPUS provides a nonhomogeneous solution when more than one station is solved
  - Use OPUS Projects, which requires training to use
- Self processing combines a single homogeneous solution
  - That is, optical and GNSS observations can be adjusted simultaneously
  - All stations are solved for simultaneously

So Why Process Your Own Data?

- Allows you to determine what went wrong and to correct actions that caused mistake
  - Will hopefully result in better field and office procedures in the future
- You can combine several different sources of data
  - Combine conventional observations with GNSS observations

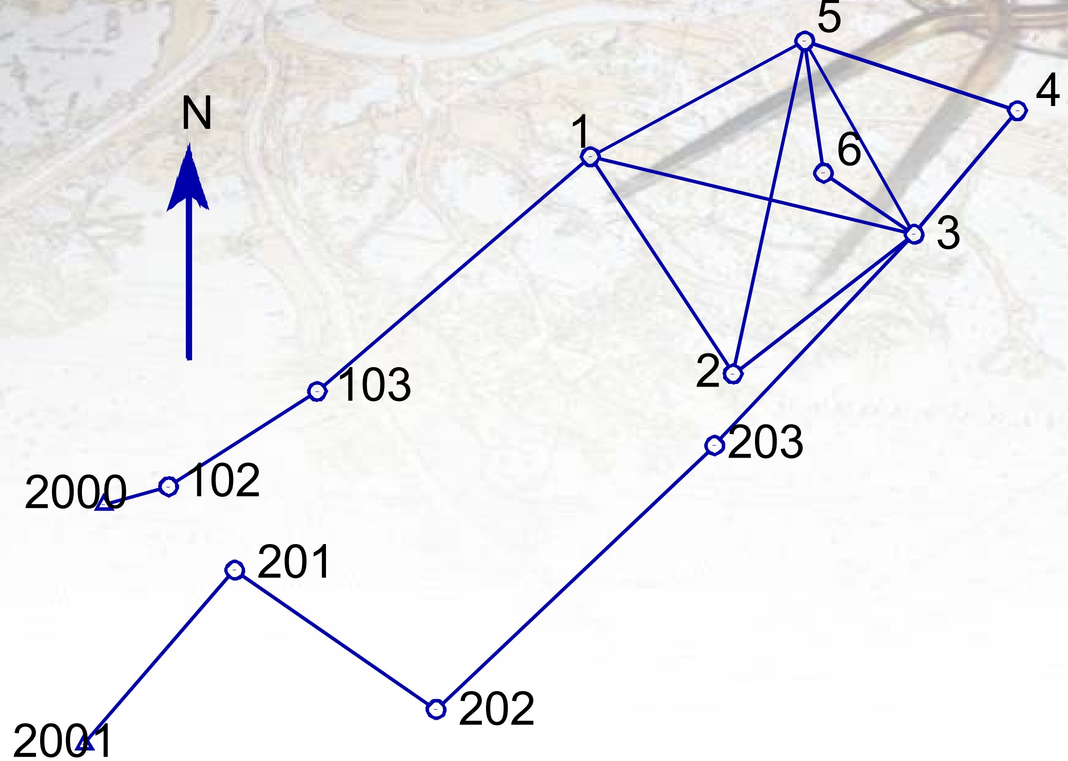
So Why Process Your Own Data?

- Use OPUS or the CORS network to connect to the national spatial reference system.
  - Create reproducible coordinates for future
  - Don’t use OPUS to create coordinates for all stations in your survey
- Professionals should know the product they deliver to a client

Procedures with Tau Criterion

1. Locate all observations that qualify for rejection
2. Reject the single observation with the largest standardized residual
3. Repeat Steps 1 and 2 until all standardized residuals are less than the rejection criterion
4. Reinsert rejected observations in a one at a time fashion to check if they should be rejected.

Example – Real from a class!



Example

- Control Stations (Second order, class II)

  2000: (419710.09, 2476334.60)

  2001: (419266.82, 2476297.98)

- Distance between them is 444.78 ft
- Distance precision: $\sigma_P = \frac{444.78}{20{,}000} = 0.022 \text{ ft}$
- Uncertainty for control coordinates: $\sigma_c = \frac{0.022}{\sqrt{2}} = \pm 0.016 \text{ ft}$

Example

Backsight	Occupied	Foresight	Angle	*S* (″)
102	2000	2001	109°10′54.0″	25.5
2000	102	103	162°58′16.0″	28.9
102	103	1	172°01′43.0″	11.8
2000	2001	201	36°04′26.2″	7.4
2001	201	202	263°54′18.7″	9.7
201	202	203	101°49′55.0″	8.1
202	203	3	176°49′10.0″	8.4
203	3	2	8°59′56.0″	6.5
2	1	3	316°48′00.5″	6.3
3	5	4	324°17′44.0″	8.1

Example

BS	Occ	FS	Angle	S (″)
6	5	3	338°36′38.5″	10.7
1	5	3	268°49′32.5″	9.8
2	5	3	318°20′54.5″	7.0
2	3	1	51°07′11.0″	7.2
2	3	5	98°09′36.5″	10.3
2	3	6	71°42′51.5″	15.1
2	3	4	167°32′28.0″	14.5

Example

- Distance Observations

From	To	Dist (ft)	S (ft)
2001	201	425.90	0.022
201	202	453.10	0.022
202	203	709.78	0.022
203	3	537.18	0.022
5	3	410.46	0.022
5	4	397.89	0.022
5	6	246.61	0.022
5	1	450.67	0.022
5	2	629.58	0.022
3	2	422.70	0.022

Example

From	To	Dist (ft)	S (ft)
3	1	615.74	0.022
3	5	410.44	0.022
3	6	201.98	0.022
3	4	298.10	0.022
1	2	480.71	0.022
1	3	615.74	0.022
2000	102	125.24	0.022
102	103	327.37	0.022
103	1	665.79	0.022

Example

- Run in ADJUST
  - Partial listing of adjusted distances

Station Occupied	Station Sighted	Distance	V	Std.Res.	Red.#
5	2	630.97	1.394	76.87	0.679
3	2	422.77	0.069	3.67	0.736
3	1	616.23	0.495	26.05	0.746
3	5	413.76	3.317	171.47	0.773
3	6	203.77	1.789	106.83	0.579
3	4	266.81	-31.287	-1802.59	0.622*
1	2	480.95	0.243	13.29	0.689
1	3	616.23	0.495	26.05	0.746
2000	102	125.05	-0.188	-28.72	0.089
102	103	327.25	-0.121	-17.16	0.103
103	1	665.70	-0.087	-12.18	0.105

Partial Listing of Adj. Angle Obs.

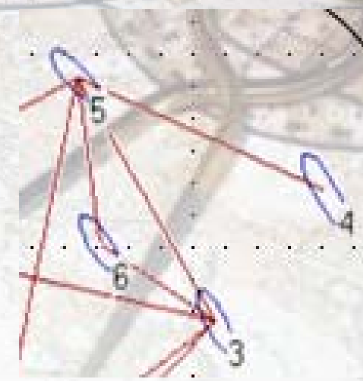
BS	Occ	FS	Angle	V	Std.Res.	Red.#
102	2000	2001	109°40'20.8"	1766.78"	106.51	0.423
2000	102	103	162°23'46.9"	-2069.09"	-110.49	0.420
102	103	1	171°57'47.3"	-235.73"	-112.47	0.032
2000	2001	201	36°07'53.9"	207.69"	104.46	0.072
2001	201	202	263°58'29.6"	250.90"	104.48	0.061
201	202	203	101°52'55.9"	180.89"	58.03	0.148
202	203	3	176°50'15.3"	65.26"	23.14	0.113
203	3	2	8°59'35.6"	-20.36"	-13.36	0.055
2	1	3	316°49'53.8"	113.33"	28.07	0.411
3	5	4	322°04'20.0"	-8003.99"	-1745.37	0.321*
6	5	3	338°42'49.1"	370.61"	62.85	0.304

From distances

3	4	266.81	-31.287	-1802.59	0.622*

Remove Largest

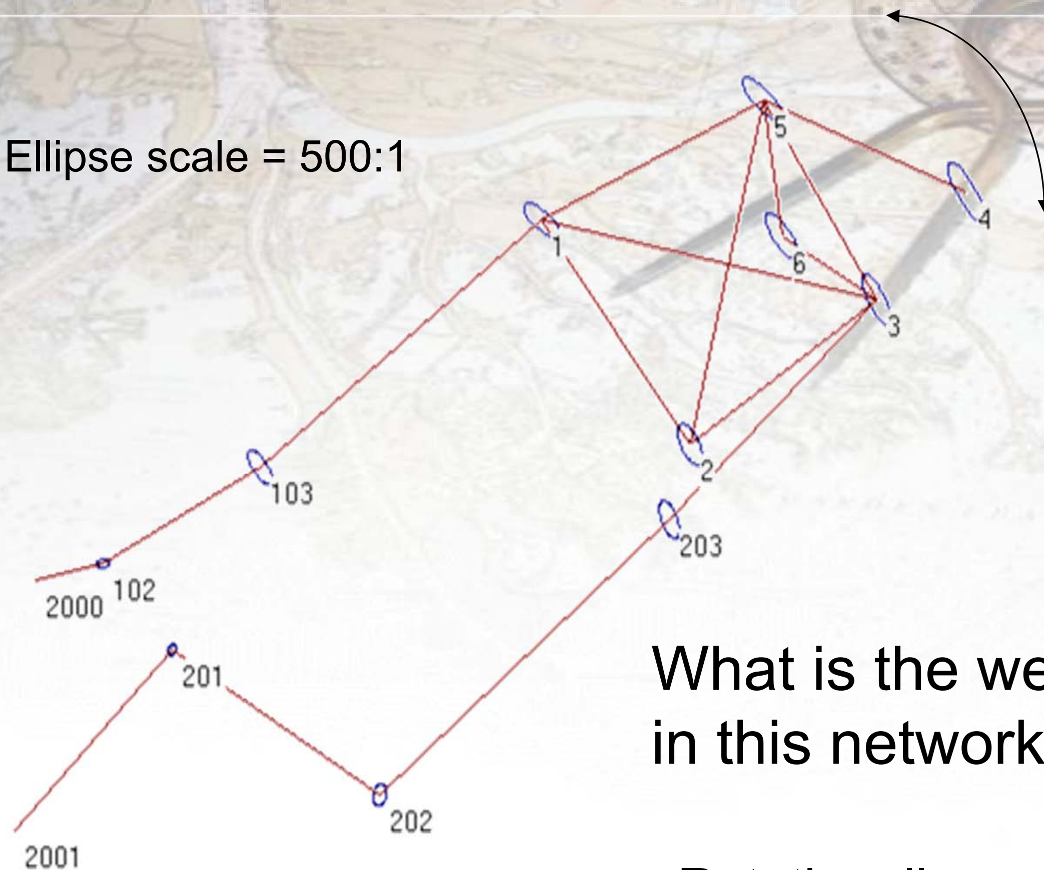
- Remove distance from 3 to 4
  - Note this distance would directly affect angle 3-5-4
- Proceed to next attempted adjustment
  - Second attempt: Remove angle 102 – 103 – 1
  - Third attempt: Success!



Example 21-1 -second attempt.adat

Plot of Error Ellipses

Ellipse scale = 500:1

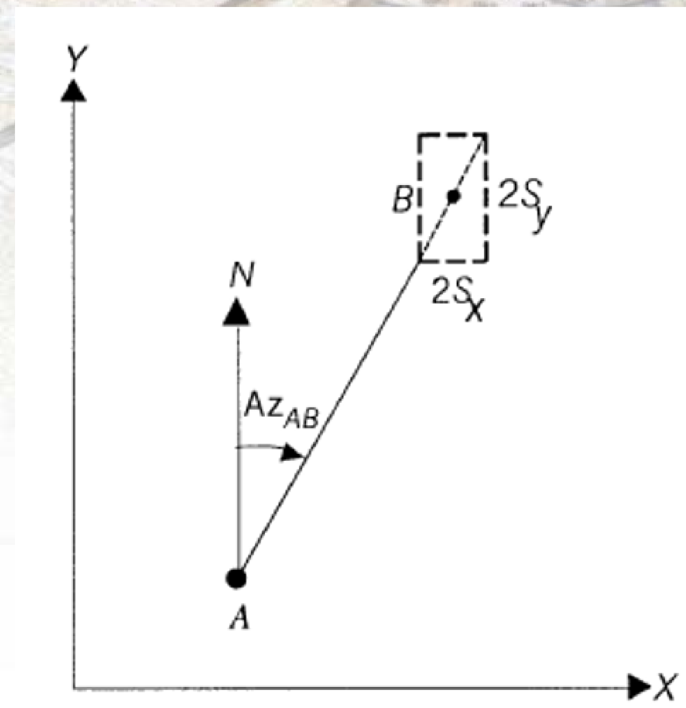


What is the weakness in this network?

Rotationally unstable!

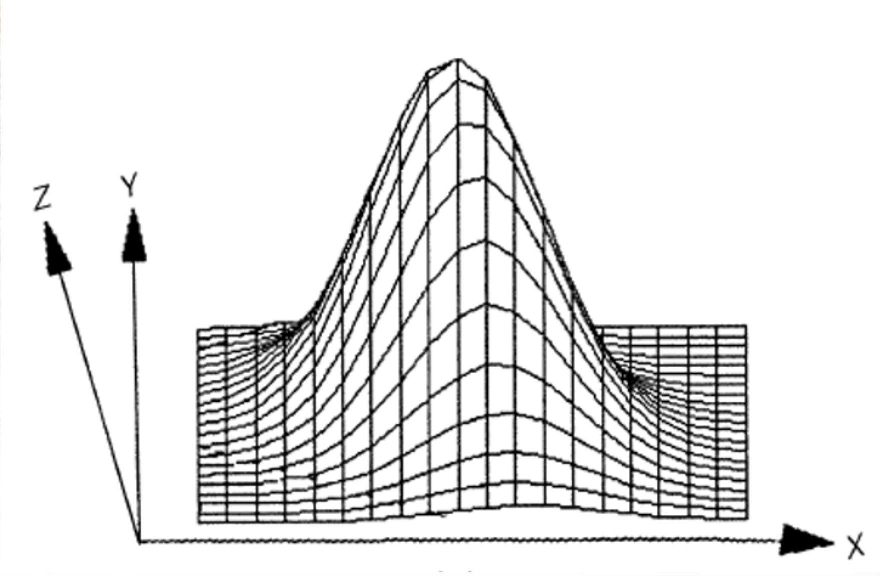
Error Ellipse

- Drawing standard deviations of coordinates about a stations creates the *standard error rectangle*



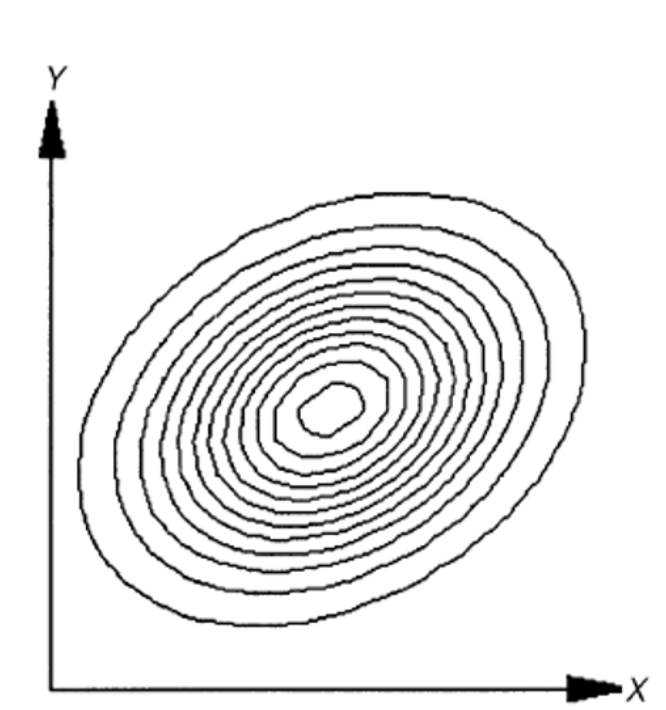
Error Ellipse

- However, at each station there is a bivariate(2 variables, which are *x* and *y*) distribution



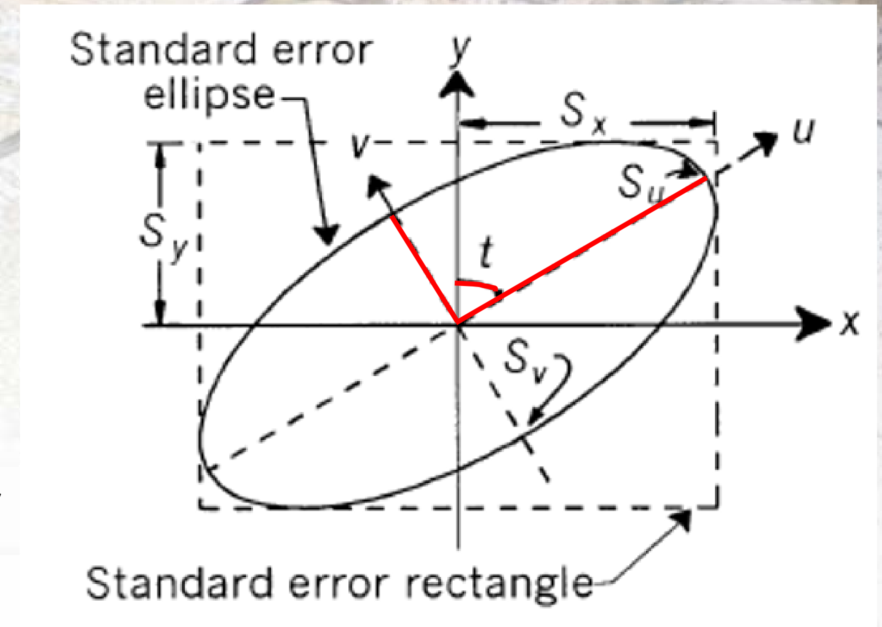
Error ellipse

- Cross sections of the bivariate distribution create error ellipses at different probability levels



Error ellipses

- Parts of an error ellipse are:
  - Minor axis, *v*
  - Major axis, *u*
  - Rotation angle, *t*, from the minor axis
  - S_v is the length of the semi-minor axis
  - S_u is the length of the semi-major axis



Computation of Error Ellipse

- Error ellipse computed from variances and covariance of unknown station coordinates
- Results in removal of correlation in station coordinates
  - Correlation implies that if one coordinate changes, the other coordinate(s) must also change.

Local Accuracy Sheet

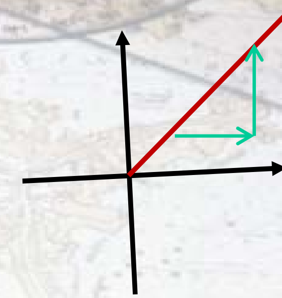
HX1559	Type/PID	Horiz	Ellip	Dist(km)	SD_N	SD_E	SD_h	CorrNE
HX1559	NETWORK	1.02	1.88		0.43	0.40	0.96	−0.20195
HX1559	LOCAL	(053 points):						
HX1559	AJ2516	1.52	2.92	16.87	0.66	0.57	1.49	−0.17057
HX1559	AJ2507	1.43	2.65	22.15	0.60	0.56	1.35	−0.19531
	⋮ and so on							
HX1559	AF9660	1.02	1.90	337.44	0.43	0.40	0.97	−0.20298
HX1559								
HX1559	MEDIAN	1.31	2.43	94.51				

Correlation between NE

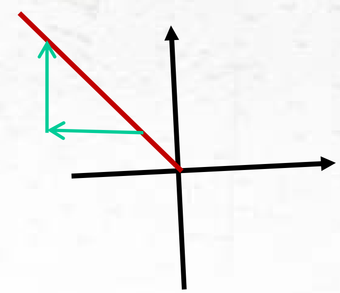
Accuracy and standard deviation values are given in cm.

What is Correlation?

- CorrNE represents correlation in Northing and Easting coordinates
  - When > 0, this implies that as N increases(decreases) so must E increase(decrease)



  - When < 0, this implies that as N increases(decreases) E must decrease(increase)



What is Correlation?

- The Pearson correlation coefficient is computed as

$$\rho(N,E) = \frac{Cov(N,E)}{\sigma(N)\sigma(E)}$$

- So for

Type	Horiz	Ellip	Dist(km)	SD_N	SD_E	SD_h	CorrNE
NETWORK	1.02	1.88		0.43	0.40	0.96	-0.20195

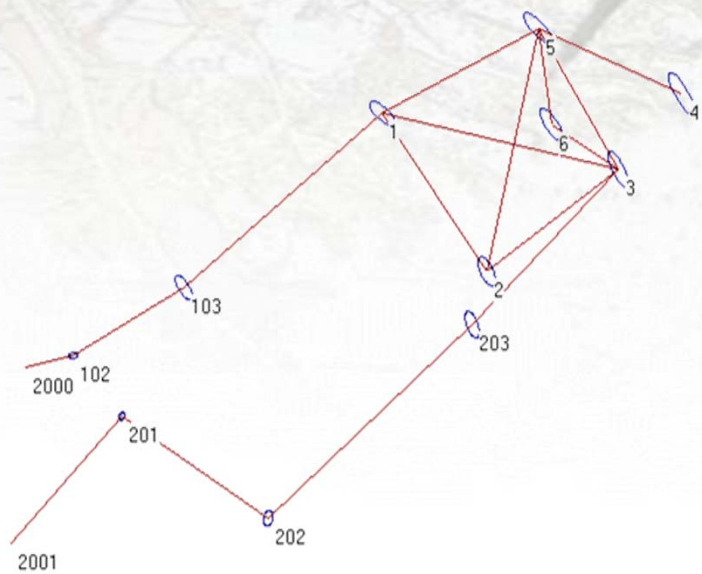
- This would mean that the Cov(N,E) would be

$$Cov(N,E) = SD_N * SD_E * \rho(N,E)$$

$$Cov(N,E) = 0.43 * 0.40 * (-0.20195) = -0.03474$$

Error Ellipses

- Error ellipses can graphically show weakness in control surveys



Radial Error at a Station

- 1998 FGDC Geospatial Positioning Accuracy Standards[a] require the computation of the radial error at a station
  - *Horizontal*: The reporting standard in the horizontal component is the radius of a circle of uncertainty, such that the true or theoretical location of the point falls within that circle 95% of the time
  - *Vertical*: The reporting standard in the vertical component is a linear uncertainty value, such that the true or theoretical location of the point falls within +/- of that linear uncertainty value 95% of the time

[a]https://www.fgdc.gov/standards/projects/accuracy/

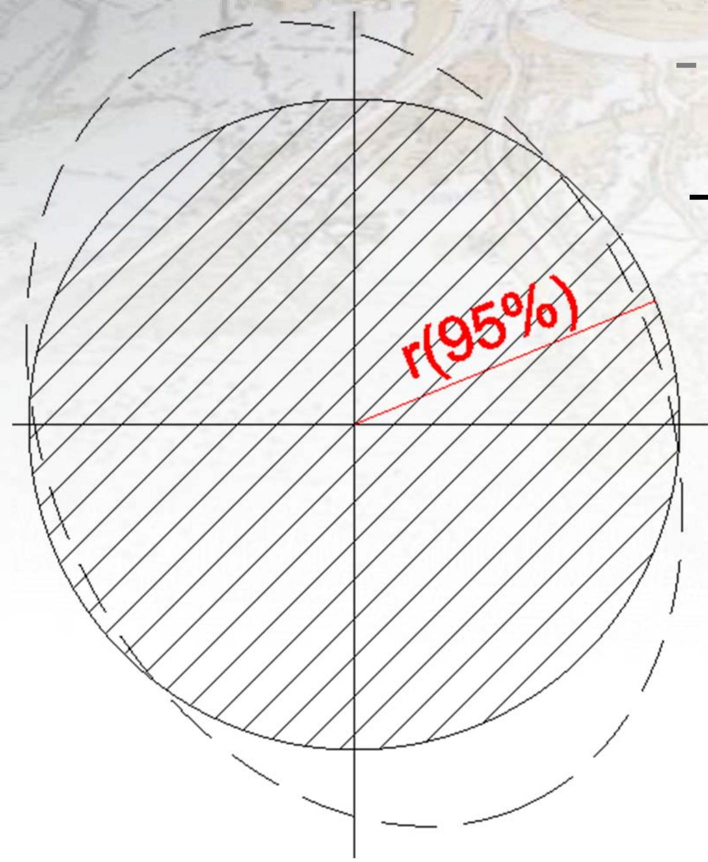
1998 FGDC Accuracy Standards

Table 2.1 Accuracy Standards: Horizontal, Ellipsoid Height, and Orthometric Height

Accuracy Classifications	95% Confidence Less Than or Equal to
1 millimeters	0.001 meters
2 millimeters	0.002 meters
5 millimeters	0.005 meters
1 centimeters	0.010 meters
2 centimeters	0.020 meters
5 centimeters	0.050 meters
1 decimeters	0.100 meters
2 decimeters	0.200 meters
5 decimeters	0.500 meters
1 meters	1.000 meters
2 meters	2.000 meters
5 meters	5.000 meters
10 meters	10.000 meters

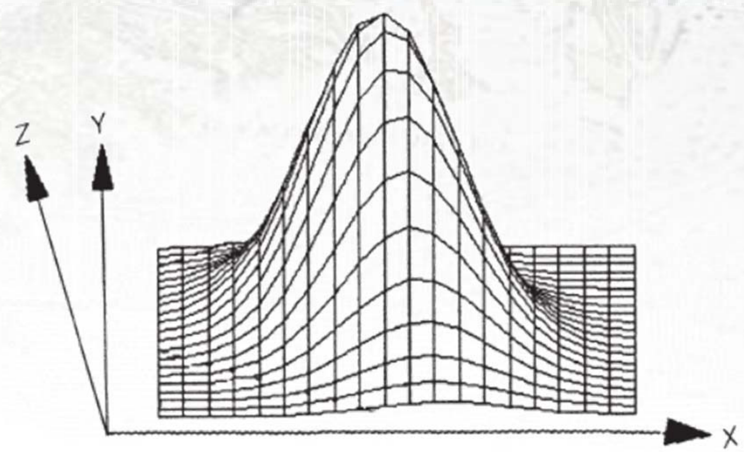
- When control points in a survey are classified, they have been verified as being consistent with all other points in the network, not merely those within that particular adjustment
- These are not observational closures within a survey
  - but the ability of that survey to duplicate already established control values.

Horizontal Adjustment
Radial Error versus Error Ellipse



- - - - 95% error ellipse

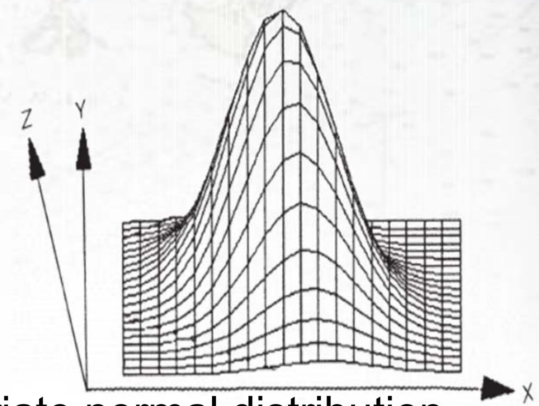
—— 95% radial error



Bivariate normal distribution

Computing Horizontal Radial Errors

- Needed at each station
  - Components of the standard error ellipse from a least squares adjustment
    - Semimajor axis of the ellipse, S_u
    - Semiminor axis of the ellipse, S_v
  - Multiplier from a bivariate normal distribution curve to report 95% radial error at station



Bivariate normal distribution

Computing Horizontal Radial Errors

- Computation of multiplier (critical value) from bivariate normal distribution is difficult
  - Approximated by Leenhouts (1985), as follows:
    - Compute $C = {}^{S_v}/_{S_u}$
    - Then 95% radial error at a station is

$$r_{95\%} = S_u\left(1.960790 + 0.004071C + 0.114276C^2 + 0.371625C^3\right)$$

- Note critical value varies with size of *C*

Example

- Assume that the error ellipse at a station has the following values. What is the radial error at the station?

  S_u = ±0.025 ft and S_v = ±0.008 ft

- $C = \frac{0.008}{0.025} = 0.320$
- The 95% radial error at the station is

$$r_{95\%} = 0.025\left(1.960790 + 0.004071C + 0.114276C^2 + 0.371625C^3\right) = 1.986$$

$$r_{95\%} = \pm 0.025(1.986) = \pm 0.05\ \text{ft}$$

What is the Root Mean Square Error?

- The NSSDA[a] uses root-mean-square error (RMSE) to estimate positional accuracy. RMSE is the square root of the average of the set of squared differences between dataset **coordinate values** and **coordinate values** from an independent source of higher accuracy for identical points.

$$RMSE = \sqrt{\frac{\sum \text{discrepancies}^2}{n}}$$

[a]NSSDA stands for National Standard for Spatial Data Accuracy

Example with a Map

	Map Coordinates			Surveyed Checkpoint Coordinates		
Point	(1) x (m)	(2) y (m)	(3) z (m)	(4) X (m)	(5) Y (m)	(6) H (m)
1	326,064.94	27,695.00	604.22	326,069.477	27,700.068	594.056
2	332,436.19	27,596.91	578.46	332,438.494	27,605.343	587.351
3	329,244.19	25,219.12	585.16	329,235.172	25,217.616	583.693
4	326,054.07	22,891.32	558.35	326,059.779	22,890.734	563.272
5	332,431.37	22,795.94	571.03	332,437.285	22,796.579	575.159

- There should at least 20 check points distributed throughout the map
- (x,y,z) are map coordinates obtained from a digital map
- (X,Y,Z) are surveyed coordinates obtained by total station survey of same points

Example using a Map

- Steps
  1. Take the difference between each map and survey coordinate
  2. Compute the *root mean square error* (RMSE) for each coordinate
- Note: These same steps would be used if network accuracy was needed for BM or horizontal survey

Computing RMSE

- Root mean square error (RMSE) should be computed as

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{[f(x_i) - x_i]^2}{n}}$$

  - Where $f(x_i)$ is the map coordinate, x_i the surveyed coordinate, and n the number of test points
  - For a horizontal or elevation survey $f(x_i)$ would be the surveyed coordinates/elevation and x_i would be the check point coordinates/elevation

Example for a Map

	Map Coordinates			Surveyed Checkpoint Coordinates			Discrepancies/Residuals		
Point	(1) x (m)	(2) y (m)	(3) z (m)	(4) X (m)	(5) Y (m)	(6) H (m)	(7) Δx	(8) Δy	(9) Δz
1	326,064.94	27,695.00	604.22	326,069.477	27,700.068	594.056	−4.537	−5.068	10.164
2	332,436.19	27,596.91	578.46	332,438.494	27,605.343	587.351	−2.304	−8.433	−8.891
3	329,244.19	25,219.12	585.16	329,235.172	25,217.616	583.693	9.018	1.504	1.467
4	326,054.07	22,891.32	558.35	326,059.779	22,890.734	563.272	−5.709	0.586	−4.922
5	332,431.37	22,795.94	571.03	332,437.285	22,796.579	575.159	−5.915	−0.639	−4.129
						Number of Checkpoints	5	5	5
						RMSE (m)	5.91	7.76	6.72
						$RMSE_r$	7.41		
						95% RMSE	12.83		13.14

Note: FGDC requires a minimum of 20 check points. Check standards!

Example for a Map

Point	Map Coordinates			Surveyed Checkpoint Coordinates			Discrepancies/Residuals		
	(1) x (m)	(2) y (m)	(3) z (m)	(4) X (m)	(5) Y (m)	(6) H (m)	(7) Δx	(8) Δy	(9) Δz
1	326,064.94	27,695.00	604.22	326,069.477	27,700.068	594.056	−4.537	−5.068	10.164
2	332,436.19	27,596.91	578.46	332,438.494	27,605.343	587.351	−2.304	−8.433	−8.891

- $\Delta x = x - X;\ \Delta y = y - Y;\ \Delta z = z - H$
- So for point 1
  - $\Delta x = 326{,}064.94 - 326{,}069.477 = -4.537\ \mathrm{m}$

Example for a Map

Point	Map Coordinates			Surveyed Checkpoint Coordinates			Discrepancies/Residuals		
	(1) x (m)	(2) y (m)	(3) z (m)	(4) X (m)	(5) Y (m)	(6) H (m)	(7) Δx	(8) Δy	(9) Δz
1	326,064.94	27,695.00	604.22	326,069.477	27,700.068	594.056	−4.537	−5.068	10.164
2	332,436.19	27,596.91	578.46	332,438.494	27,605.343	587.351	−2.304	−8.433	−8.891
3	329,244.19	25,219.12	585.16	329,235.172	25,217.616	583.693	9.018	1.504	1.467
4	326,054.07	22,891.32	558.35	326,059.779	22,890.734	563.272	−5.709	0.586	−4.922
5	332,431.37	22,795.94	571.03	332,437.285	22,796.579	575.159	−5.915	−0.639	−4.129

- To Compute RMSE for Δx

$$RMSE_{\Delta x} = \sqrt{\frac{4.537^2 + 2.304^2 + 9.018^2 + 5.709^2 + 5.915^2}{5}} = 5.91$$

Computing Radial Error

	(7) Δx	(8) Δy	(9) Δz
RMSE (m)	5.91	7.76	6.72
$RMSE_r$	7.41		
95% RMSE	12.83		13.14

- The radial error is then computed as

$$RMSE_r = \sqrt{RMSE_x^2 + RMSE_y^2}$$

$$RMSE_r = \sqrt{5.91^2 + 4.77^2} = 7.41$$

Compute 95% Radial Error

- From bivariate normal distribution, the 95% radial error use the formula

$$RMSE_{r@95\%} = \frac{2.4477}{\sqrt{2}} RMSE_r = 1.7308\, RMSE_r$$

$$RMSE_{r@95\%} = 1.7308\,(7.41) = \pm 12.83$$

- To compute the 95% vertical error
  - From normal distribution, multiplier is simply 1.96

$$RMSE_{H@95\%} = 1.96(6.72) = \pm 13.14$$

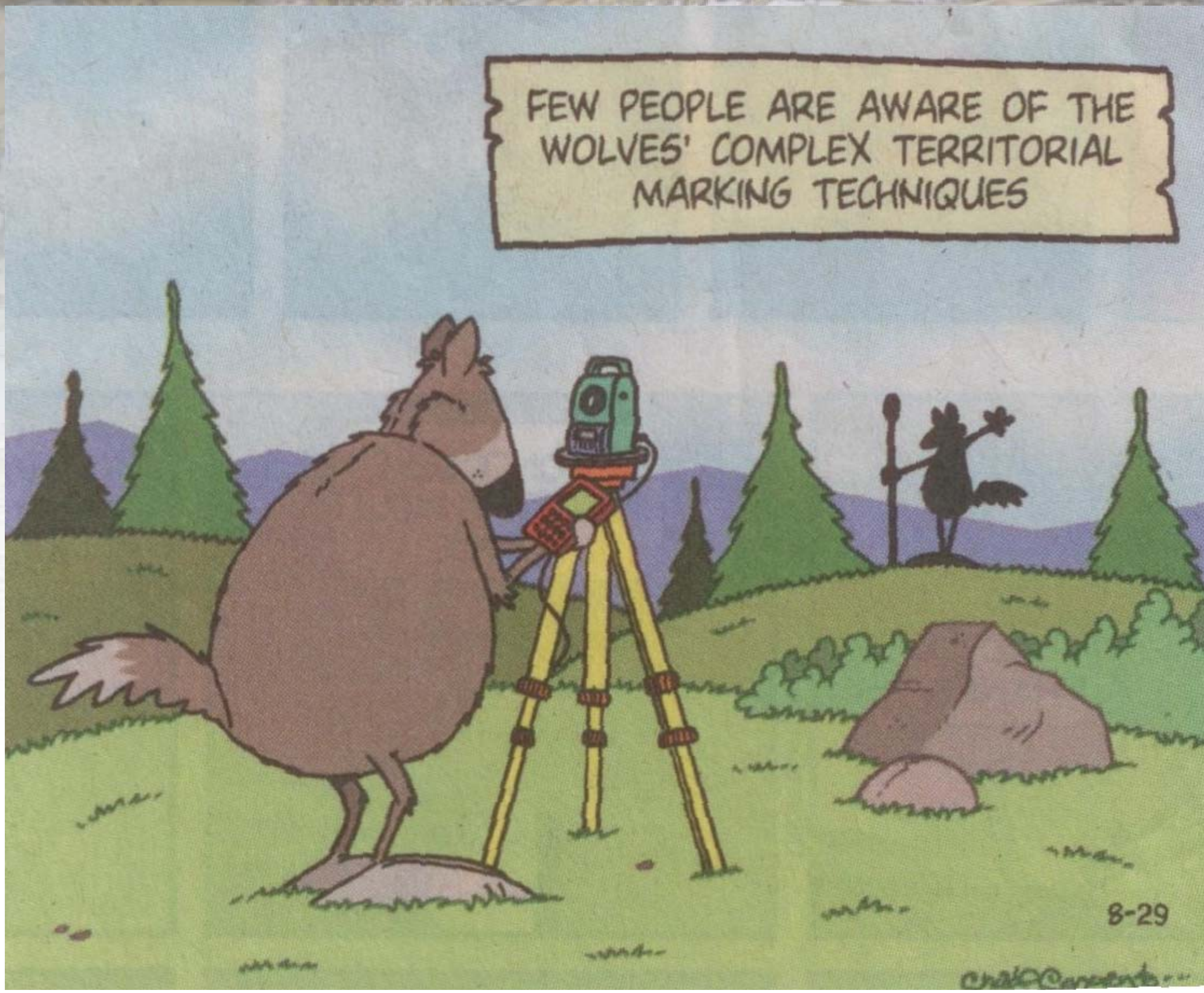
Vertical Network Accuracy

- Vertical network work accuracy computed similarly
- That is compute the RMSE for the elevations

$$\mathrm{RMSE}_z = \sqrt{\frac{(10.164)^2+(-8.891)^2+1.467^2+(-4.922)^2+(-4.129)^2}{5}} = \pm 6.72\ \mathrm{m}$$

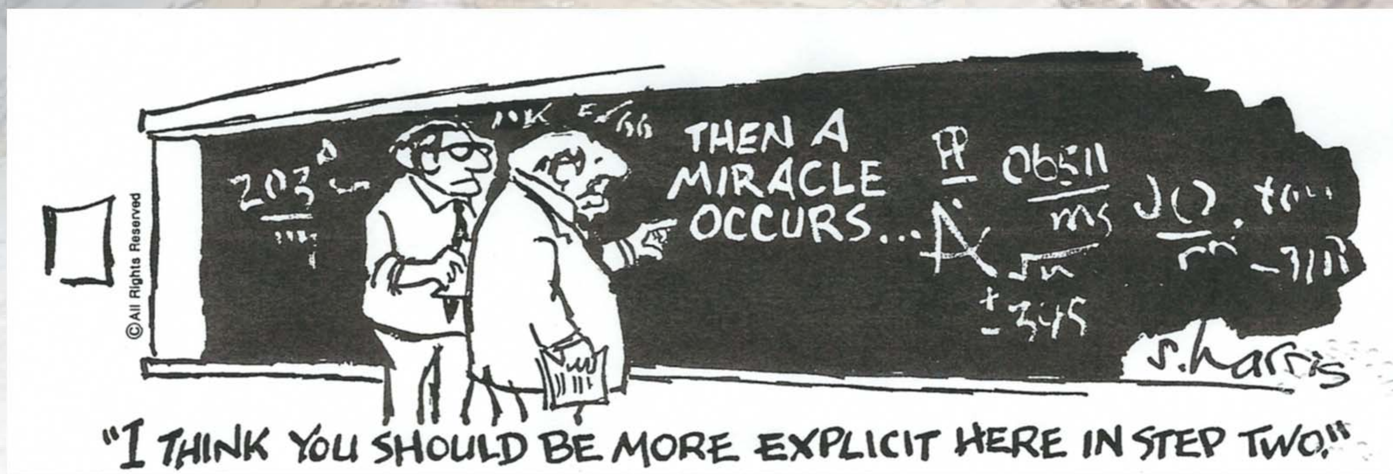
- The 95% network accuracy is

$$RMSE_{z@95\%} = 1.96(RMSE_z) = 1.96(6.72) = 13.2\ \mathrm{m}$$


FEW PEOPLE ARE AWARE OF THE WOLVES' COMPLEX TERRITORIAL MARKING TECHNIQUES
8-29

If a man empties his purse into his head, no man can take it away from him. An investment in knowledge always pays the best interest.

- Benjamin Franklin


THEN A MIRACLE OCCURS...
"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."
©All Rights Reserved
S. Harris

Questions?