Numerical Linear Algebra Before the Last Paradigm Shift*

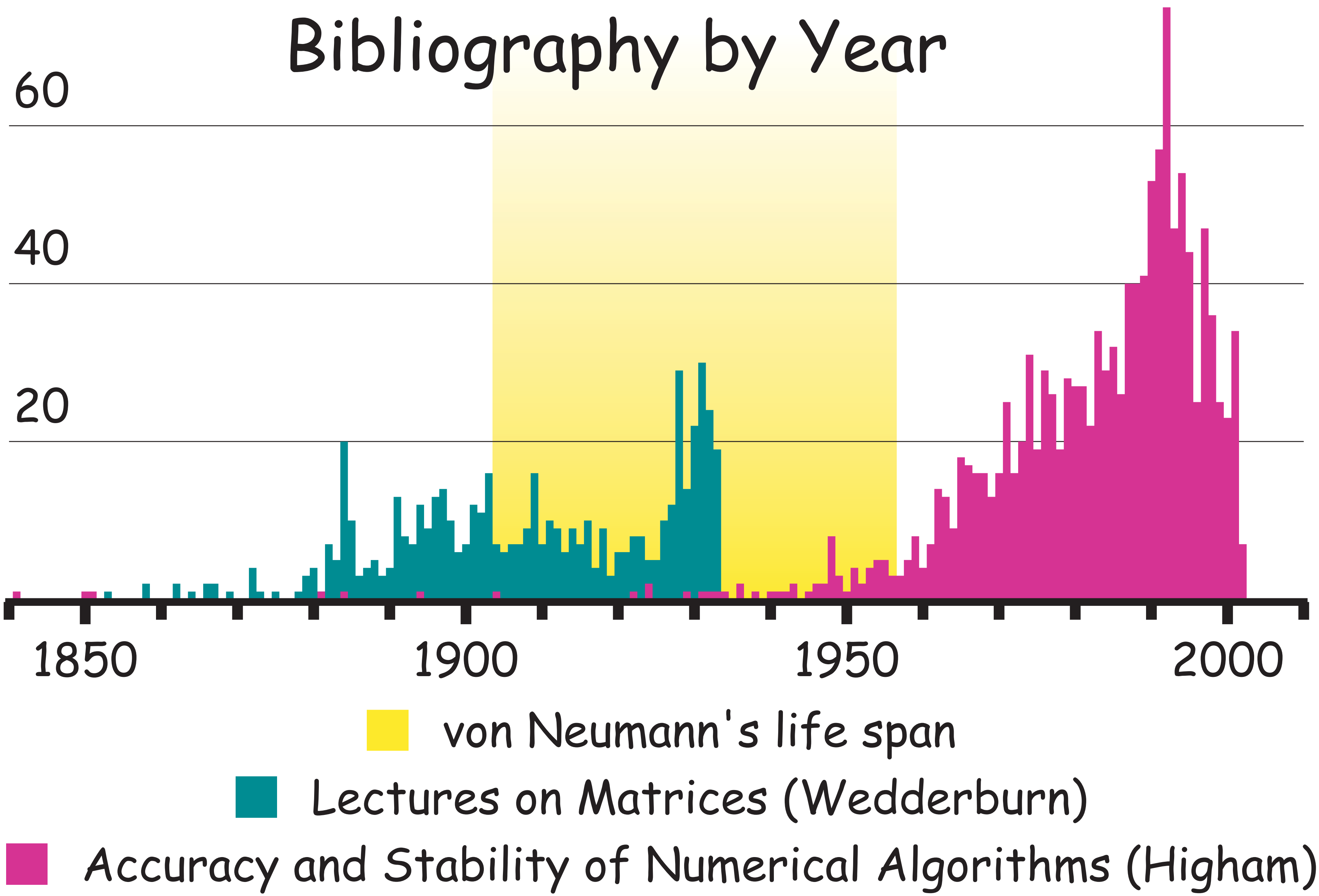
Joseph F. Grcar
Lawrence Berkeley Laboratory
jfgrcar@lbl.gov

October 5, 2005
Berkeley

*Thomas Kuhn, *The Structure of Scientific Revolutions*, The University of Chicago Press, Chicago, 1970.

“The state of numerical mathematics stayed pretty much the same as Gauss left it until World War II.”

— Herman Goldstine, 1972

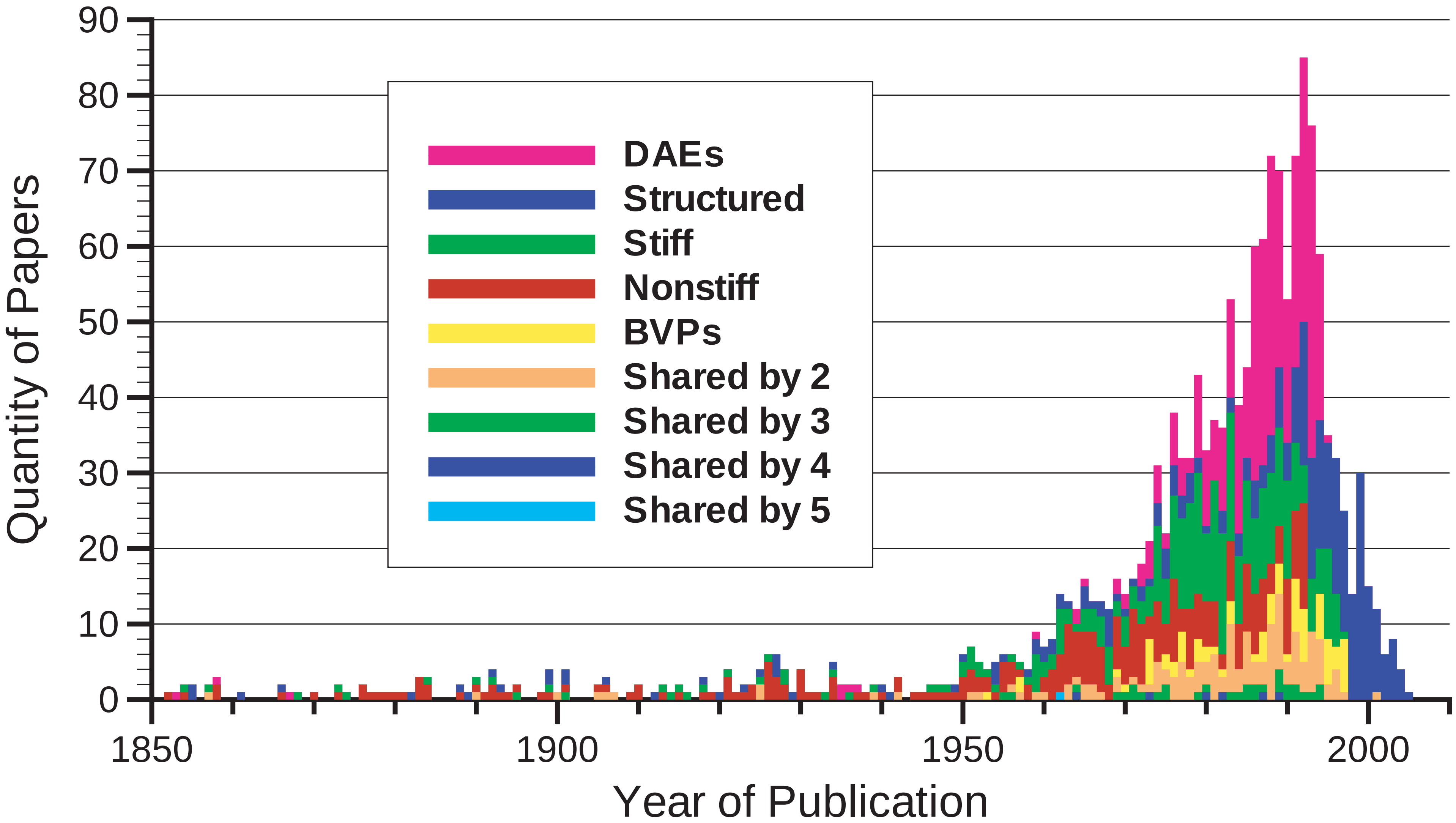


"The state of numerical mathematics stayed pretty much the same as Gauss left it until World War II."

— Herman Goldstine, 1972

Size of Numerical ODE Literature

Bibliographies of Hairer, Petzold, Wanner, Werner, et al.



Outline

Frobenius norm

matrices

adjustment of observations

survey adjustments

Gaussian elimination

geodesy

matrix factoring

Frobenius norm

$$||\boldsymbol{A}||_{\boldsymbol{F}}$$

Ferdinand Frobenius

1849 – 1917

Sitzungsberichte der Koniglich Prußischen Akademie der Wissenschaften zu Berlin, 241 – 248 (1911)

Let $r_1, \ldots, r_n$ be the characteristic roots of a matrix of the form

$$R = \sum r_{i,j} x_i y_j$$

I define the *Spur*

$$\chi(R) = \sum r_{i,i} = \sum r_i$$

and the *Spannung*

$$\vartheta(R) = \chi(R\,\bar{R}') = \|R\|_F^2$$

Ferdinand Frobenius

1849 – 1917

Sitzungsberichte der Koniglich Prußischen Akademie der Wissenschaften zu Berlin, 241 – 248, 654–665 (1911)

The *Spannung*

$$\vartheta(R) = \|R\|_F^2$$

obeys the Schwartz inequality

$$\sqrt{\vartheta(P-Q)} \leq \sqrt{\vartheta(P)} + \sqrt{\vartheta(Q)}$$

and also

$$\vartheta(P\,Q) \leq \vartheta(P)\,\vartheta(Q)$$

John von Neumann
1903 – 1957

Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren, *Mathematische Annalen*, 102:49–131 (1929)

Von Neumann was mostly interested in unbounded operators.

1932 Banach's Banach spaces

Stefan Banach

1892 – 1945

Théorie des Opérations linéaires, Warsaw (1932)

The *norme* $|U|$ of a linear operator $U(x)$ is the smallest number M satisfying

$$|U(x)| \leq M\,|x|$$

for all x

Joseph Wedderburn
1892 – 1945

Lectures on Matrices, American Mathematical Society, New York (1934)

His extensive bibliography is the source for references to Peano, Schur and the earlier Frobenius paper.

He chose the *absolute value*

$$|\cdot| = \|\cdot\|_F$$

1937 von Neumann, again

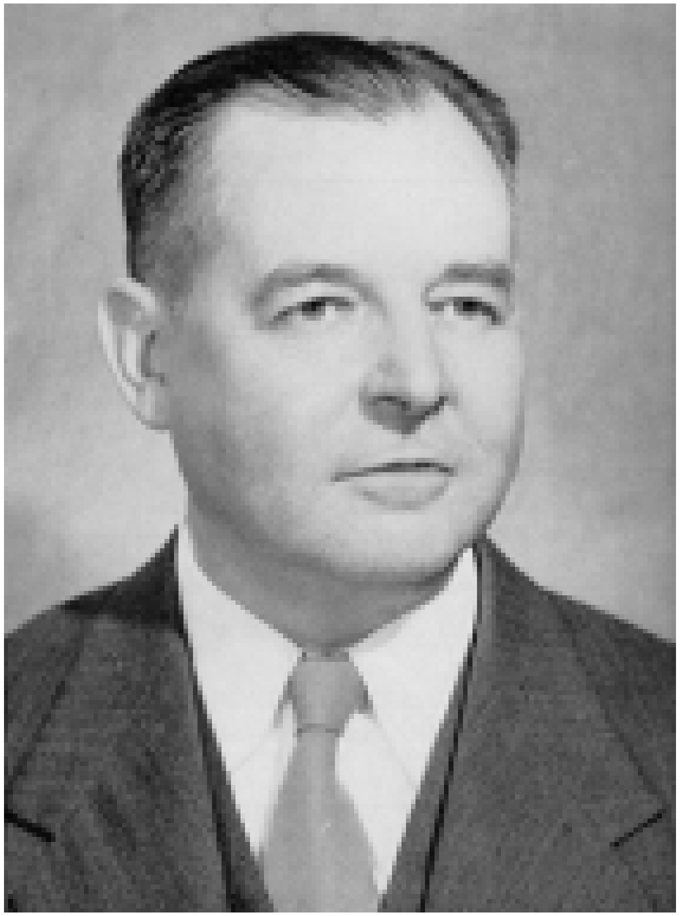
John von Neumann
1903 – 1957

Some Matrix-Inequalities and Metrization of Matric-Space, *Tomsk Univ. Rev.* (1937)

Let φ be a symmetric gauge function.

For a matrix X, form XX^* and its proper values $\omega_1, \ldots,$ ω_n, and define

$$|X|_\varphi = \varphi(\omega_1^{1/2}, \ldots, \omega_n^{1/2})$$



Harold Hotelling

1895 – 1973

Some New Methods in Matrix Calculation, *The Annals of Mathematical Statistics*, 14:1–34 (1943)

“The *norm* of the matrix $\boldsymbol{A}$

$$\boldsymbol{N(A)} = \|\boldsymbol{A}\|_{\boldsymbol{F}}$$

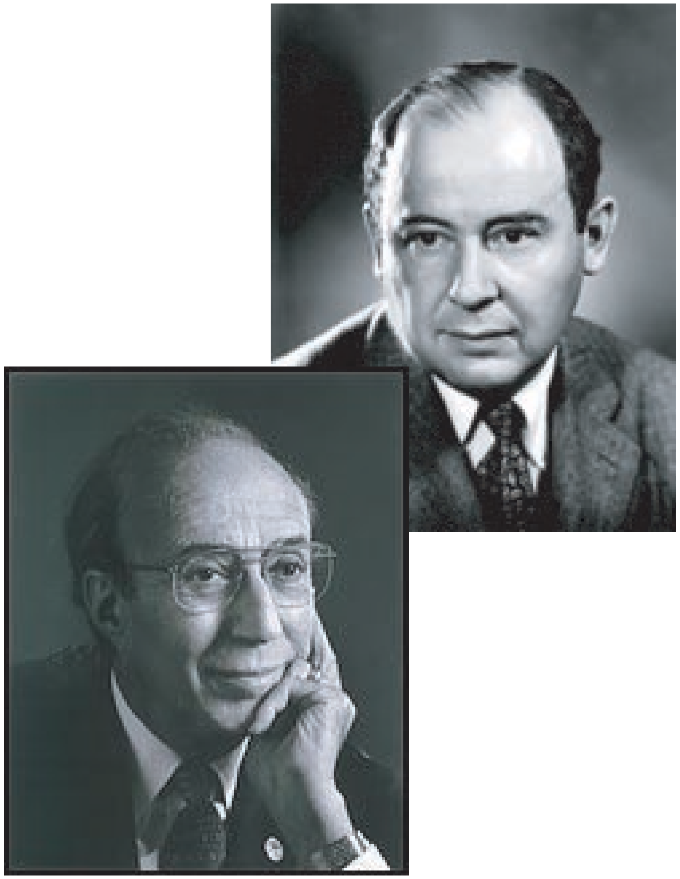
is what Wedderburn defines as the absolute value.”

Cyrus MacDuffee
1895 – 1961

The Theory of Matrices, New York (1946)

“Notation for matrices is far from agreed.”

1947 von Neumann and Goldstine



von Neumann 1903 – 1957
and Goldstine 1913 – 2004

Numerical Inverting of Matrices of High Order, *Bulletin of the American Mathematical Society*, 53:1021–1099 (1947)

norm

$$\boldsymbol{N(A)} = \|\boldsymbol{A}\|_{\boldsymbol{F}}$$

upper bound

$$|\boldsymbol{A}| = \max_{|\boldsymbol{x}|=1} |\boldsymbol{Ax}| = \|\boldsymbol{A}\|_{\boldsymbol{2}}$$

lower bound

$$|\boldsymbol{A}|_{\ell} = \min_{|\boldsymbol{x}|=1} |\boldsymbol{Ax}|$$

Alan Turing
1912 – 1954

Rounding-off Errors in Matrix Processes, *The Quarterly Journal of Mechanics and Applied Mathematics*, 1:287–308 (1948)

norm

$$\boldsymbol{N(A) = \|A\|_F}$$

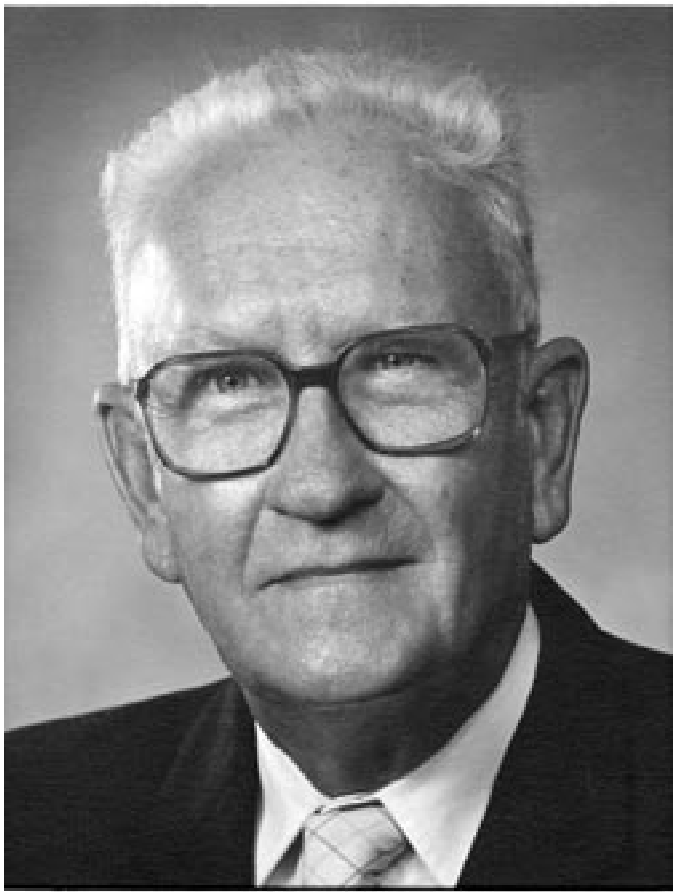
maximum expansion

$$\boldsymbol{B(A) = \|A\|_2}$$

maximum coefficient

$$\boldsymbol{M(A) = \max_{i,j} |a_{i,j}|}$$

1953 Householder



Alston Householder

1904 – 1993

Principles of Numerical Analysis, McGraw-Hill, New York (1953)

norm

$$N(A) = \|A\|_F$$

maximum

$$M(A) = \|A\|_2$$

bound

$$b(A) = \max_{i,j} |a_{i,j}|$$

1956, 1959 Bodewig

E(wald?) Bodewig

Matrix Calculus, North-Holland, Amsterdam (1956, 1959)

norm

$$N(A) = \|A\|_F$$

maximum expansion or bound

$$|A| = \|A\|_2$$

maximum coefficient

$$m(A) = \max_{i,j} |a_{i,j}|$$

1963 Wilkinson

James Wilkinson

1919 – 1986

Rounding Errors in Algebraic Processes, Prentice Hall, Amsterdam (1963)

Euclidean or *Schur norm*

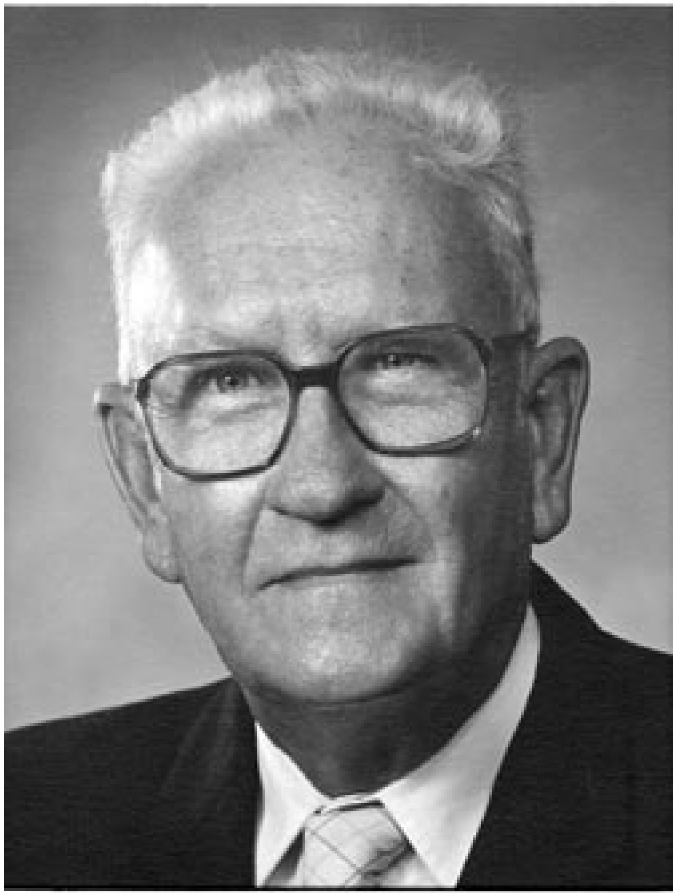
$$\|A\|_E = \|A\|_F$$

Question:

Who introduced $\|\cdot\|$ notation?

Possible Answer:

Maybe Bauer or Stoer in the 1960's.



Alston Householder

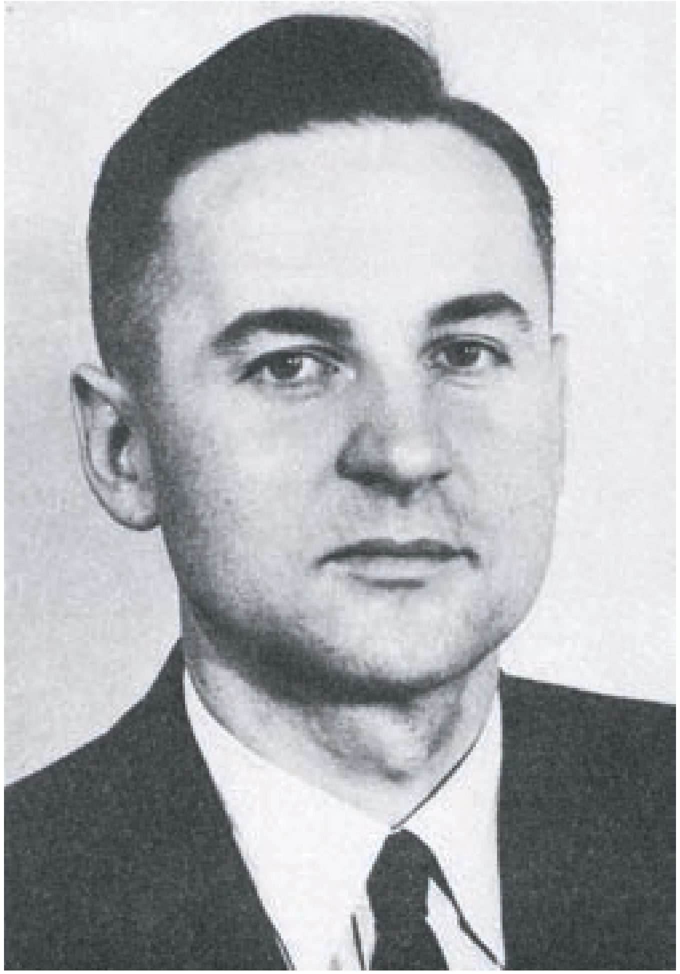
1904 – 1993

The Theory of Matrices in Numerical Analysis, Blaisdell, New York (1964)

Euclidean length of a matrix

$$\tau^{1/2}(A^H A) = \|A\|_F$$

1970 Rutishauser



Heinz Rutishauser
1918 – 1970

Lectures on Numerical Mathematics, Birkhäuser, Boston (1990), posthumous

the so-called *Schur norm*

$$\|A\|_2 = \|A\|_F$$

1973 Stewart

Pete Stewart

Introduction to Matrix Computations, Academic Press, New York (1973)

the *Frobenius norm*

$$\|\boldsymbol{A}\|_{\boldsymbol{F}}$$

Historical Pitfalls, Illustrated by $\|\cdot\|_F$

1. Anachronism:

   “Historical description means understanding things as people understood them *then*.”

2. Chronology:

   “Chronology is not the same thing as history. History is processed chronological data.”

Kenneth May (1915 – 1977)

“Historiography: A Perspective for Computer Scientists,” in *A History of Computing in the 20th Century*, edited by Metropolis, Howlett, Rota, 1980.

$\|\cdot\|_F$ Chronology

1887	Peano	used value of $\|\cdot\|_F$
1909	Schur	used value of $\|\cdot\|_F$
1911	Frobenius	*Spannung* $\vartheta(\cdot) = \|\cdot\|_F^2$ with sum and product rules
1929	von Neumann	Hilbert spaces
1932	Banach	Banach space operator $\|\cdot\|$
1934	Wedderburn	*absolute value* $\mathbf{\|} \cdot \mathbf{\|} = \|\cdot\|_F$
1937	von Neumann	matrix space metrization $\|\cdot\|_\varphi$
1943	Hotelling	*norm* $N(\cdot) = \|\cdot\|_F$
1946	MacDuffee	“notation is far from agreed”

$\| \cdot \|_F$ Chronology, continued

1947	von Neumann and Goldstine	*norm* $N(\cdot) = \| \cdot \|_F$ and *upper bound* $\lvert \cdot \rvert = \| \cdot \|_2$
1948	Turing	$N(\cdot)$, $B(\cdot)$, $M(\cdot)$
1953	Householder	$N(\cdot)$, $M(\cdot)$, $b(\cdot)$
1954	Givens	$N(\cdot)$
1959	Bodewig	$N(\cdot)$, $\lvert \cdot \rvert$, $m(\cdot)$
1963	Wilkinson	*Euclidean, Schur norm* $\| \cdot \|_E$
1964	Householder	*Euclidean length* $\tau^{1/2}(\cdot^H \cdot)$
1970	Rutishauser	*Schur norm* $\| \cdot \|_2 = \| \cdot \|_F$
1973	Stewart	*Frobenius norm* $\| \cdot \|_F$

$\|\cdot\|_F$ History

1. Value $\|\cdot\|_F$ was used by early matrix researchers, and by Frobenius especially. The real story is how it came to be named after Frobenius.

2. Hotelling deliberately ignored Wedderburn and chose *norm* $N(\cdot)$; perhaps to avoid Germanisms?

3. Von Neumann and Goldstine knew the proper notation (operator norms existed in functional analysis) but they deferred to Hotelling. Their choice led others to adopt $N(\cdot)$ for many years.

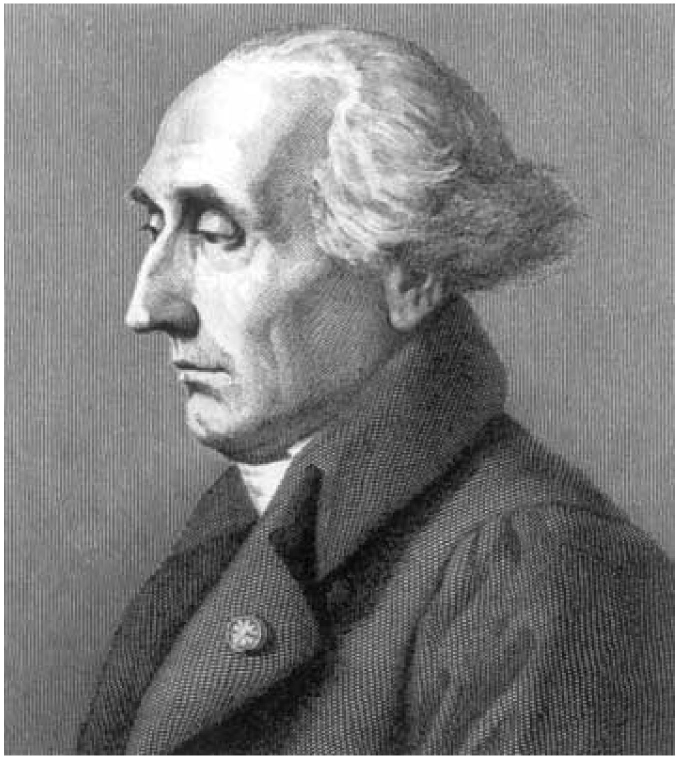
4. Attempt to use *Schur norm* $\|\cdot\|_E$ was supplanted by *Frobenius norm* $\|\cdot\|_F$ in the 1970's.

$\|\cdot\|_F$ Pertinent to the Inversion Paper

- Influence of the Inversion Paper can be seen in the number of people who adopted $N(\cdot)$ because von Neumann did.
- Why did von Neumann defer to Hotelling in using this eccentric notation?

Matrices

1759 Lagrange



Joseph-Louis Lagrange

1736 – 1813

Quadratic form of x_1, x_2, x_3

$$f = 5x_1^2 + 20x_1x_2 + 19x_2^2 + 10x_1x_3 + 22x_2x_3 + 6x_3^2$$

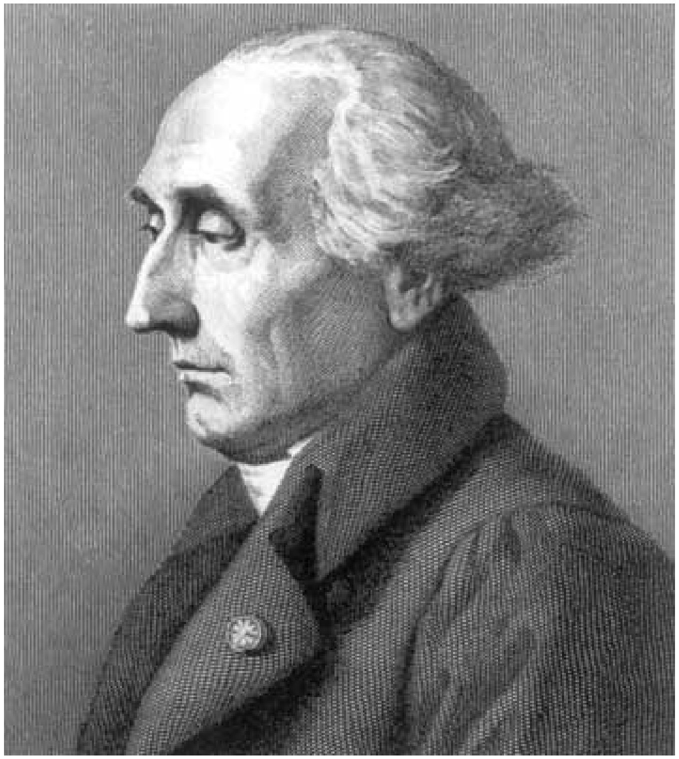
Substitute

$$x_1 = {\color{red}u_1} - \frac{1}{2}\left(\frac{20}{5}x_2 + \frac{10}{5}x_3\right)$$

to get

$$f = 5{\color{red}u_1}^2 - x_2^2 + 2x_2x_3 + x_3^2$$

1759 Lagrange



Joseph-Louis Lagrange

1736 – 1813

Quadratic form is

$$f = x^t A x$$

where $A = U^t D U$. *Linear substitution* $u = Ux$ gives

$$f = u^t D u$$

Lagrange used the weighted *sum of squares* to identify local extrema.

J. C. F. Gauss

1777 – 1855

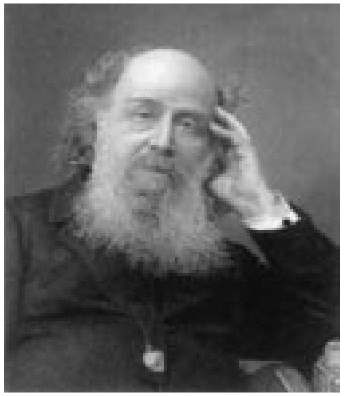
Disquisitiones arithmeticae, Leipzig, art. no. 266 (1801)

Studied quadratic forms of 2 and 3 variables.

"It is sufficient to draw this broad field to the attention of geometers.* *There is ample material for the exercise of their genius*."

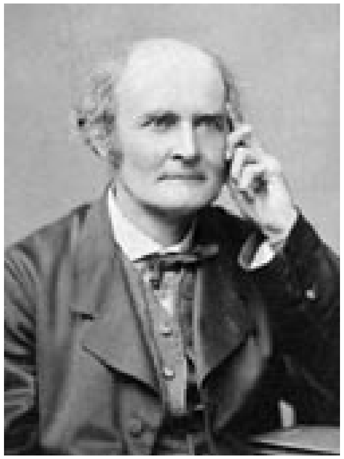
* geometer = mathematician

1852 – 1882 Matrix Algebra



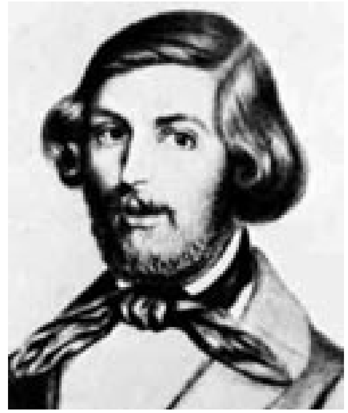
Sylvester

1814 – 1897



Cayley

1821 – 1895



Eisenstein

1823 – 1852

Laguerre

1834 – 1886

Frobenius

1849 – 1917

Invented by several people more or less independently to represent the linear substitutions that were used to study quadratic and bilinear forms.

T. Hawkins, “Another look at Cayley and the theory of matrices,” *Archives Internationales d’Histoire des Sciences*, 27(100):82–112 (1977).

1857, 1907 Not Cholesky's Factorization

Toeplitz 1881 – 1940

Jacobi 1804 – 1851

Toeplitz, "Die Jacobische Transformation der quadratischen Formen von unendlichvielen Veränderlichen," *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, 101–109 (1907).

Restated Jacobi's result "known to Lagrange and Gauss" using "matrix calculus of Frobenius."

A quadratic form

$$S = \sum_{i,k=1}^{n} \alpha_{i,k} x_i x_k$$

could be expressed as

$$S^{-1} = U'U$$

or $S = U^{-1}U'^{-1}$

Not Cholesky's Factorization

Toeplitz 1881 – 1940

Jacobi 1804 – 1851

Toeplitz, "Die Jacobische Transformation der quadratischen Formen von unendlichvielen Veränderlichen," *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, 101–109 (1907).

Restated Jacobi's result "known to Lagrange and Gauss" using "matrix calculus of Frobenius."

A quadratic form

$$S \quad n$$

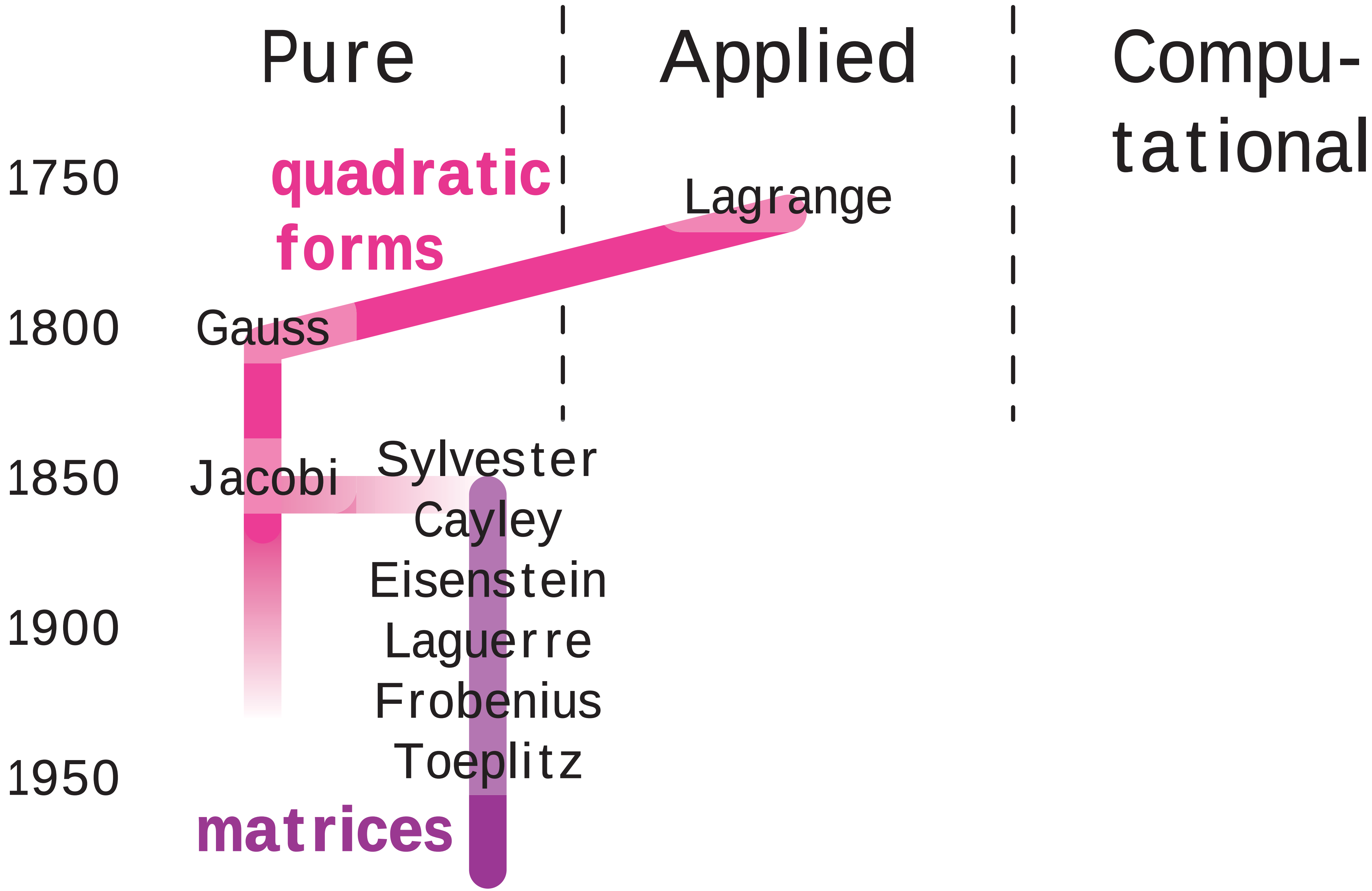
Compuationally Useless Formula of Determinants

could be expressed as

$$S^{-1} = U'U$$

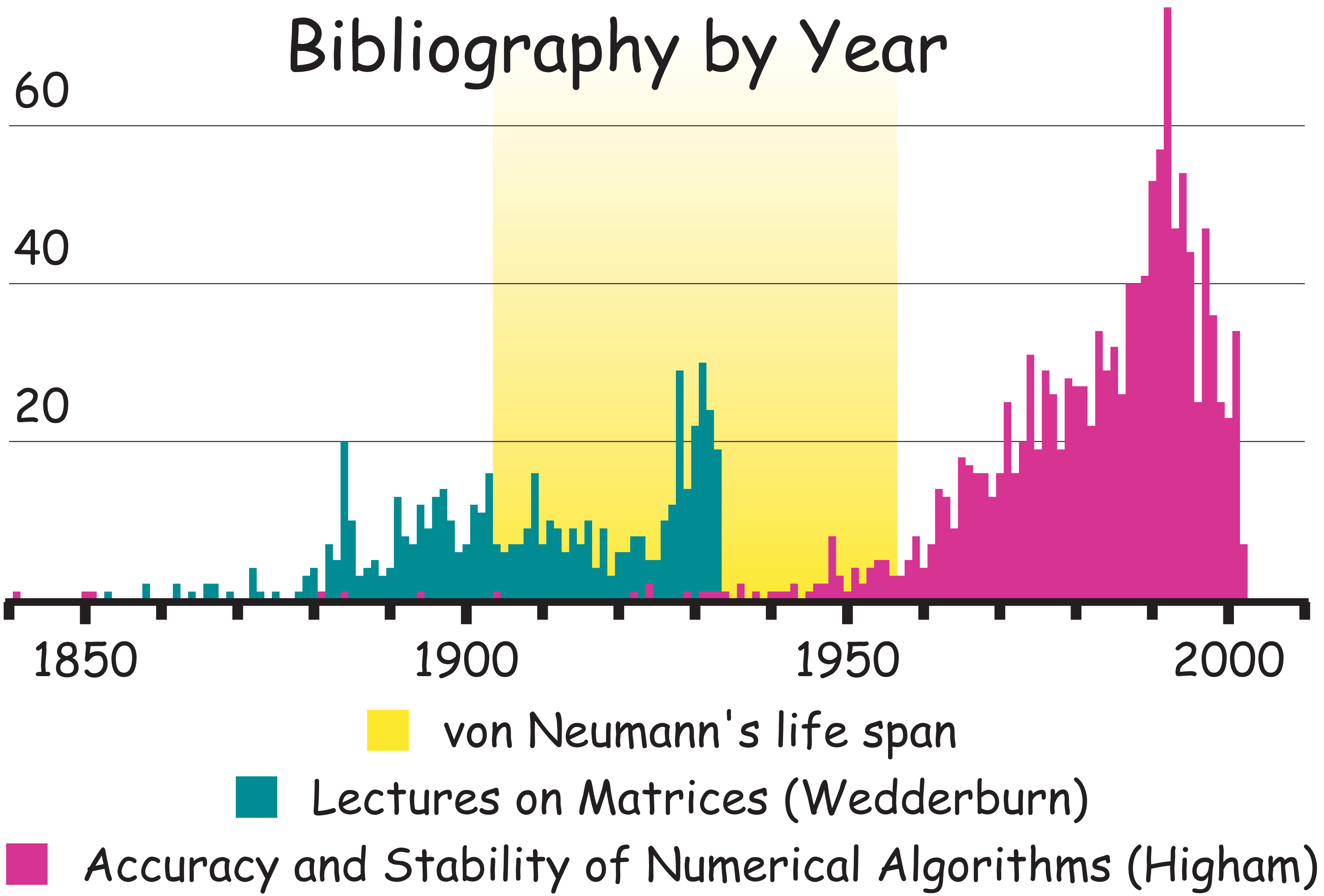
or $S = U^{-1}U'^{-1}$

Linear Algebra Timeline



“The state of numerical mathematics stayed pretty much the same as Gauss left it until World War II.”

— Herman Goldstine, 1972



Adjustment of Observations

Errors of Observation

Adrien-Marie Legendre

1752 – 1833

Johann Carl Friedrich Gauss

1777 – 1855

Pierre-Simon Laplace

1749 – 1827

Statistical theory showed

$$\left.\begin{array}{r}\text{optimal}\\ \text{estimation}\\ \text{problems}\end{array}\right\} \equiv \left\{\begin{array}{l}\text{minimizing}\\ \text{quadratic}\\ \text{functions}\end{array}\right.$$

Legendre: la méthode des moindres quarrés

Nouvelle méthodes pour la détermination des orbites des comètes, Paris (1805)

Gauss: developed probabilistic justification

Theoria Motus Corporum Coelestium in Sectionibus Conicus Solem Ambientium, Hamburg (1809)

Observations Needing Adjustment

case 1. Over-determined observations

$$\min_{\boldsymbol{x}} \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2$$

solved by $\boldsymbol{A}^t\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$ with $\boldsymbol{c} = \boldsymbol{A}^t\boldsymbol{b}$

case 2. Under-determined observations

$$\min_{\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}} \|\boldsymbol{x}\|_2$$

solved by $\boldsymbol{A}\boldsymbol{A}^t\boldsymbol{u} = \boldsymbol{b}$ where $\boldsymbol{x} = \boldsymbol{A}^t\boldsymbol{u}$

Both resulted in *normal equations*

Observations Needing Adjustment

case 1. Over-determined observations

$$\min_{\boldsymbol{x}} \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2$$

found in Astronomy

case 2. Under-determined observations

$$\min_{\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}} \|\boldsymbol{x}\|_2$$

found in Geodesy

Both resulted in *normal equations*

Aside: Astronomical Calculations

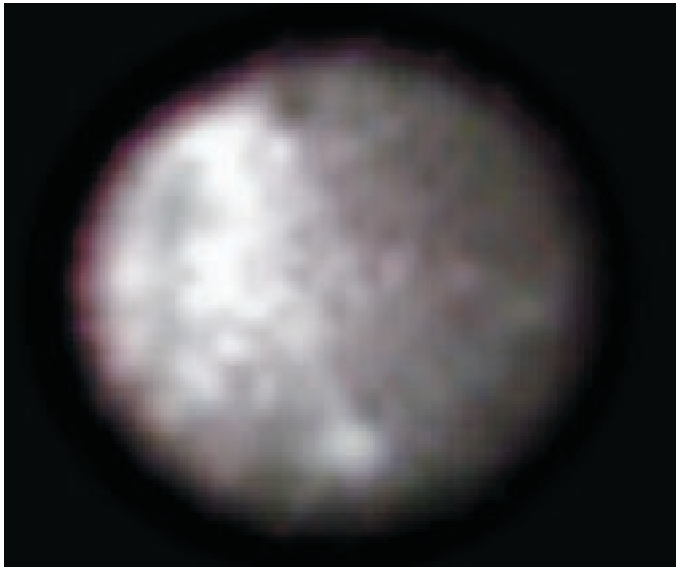
Trophy Calculations:

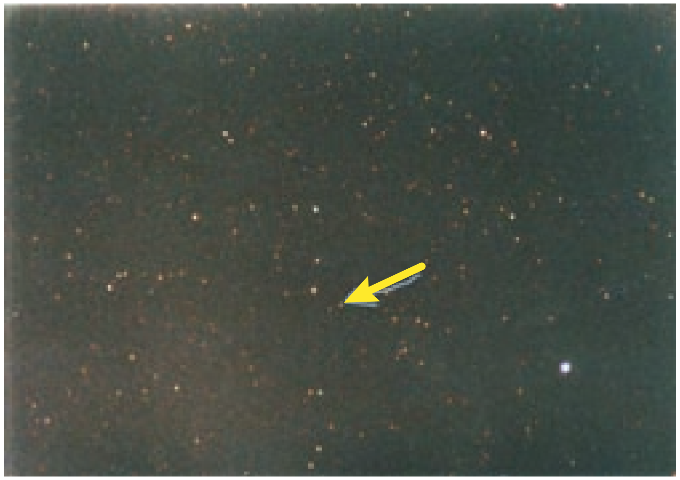
1. Fitting orbit of Ceres (over-determined observations) — Gauss

2. Triangulating distance to Cygni — Bessel

3. . . .

Nievergelt, “A tutorial history of least squares with applications to astronomy and geodesy,” in *Numerical Analysis: Historical Developments in the 20th Century*, edited by Brezinski and Wuytack, North Holland, 2001.





Aside: Astronomical Calculations

Trophy Calculations:

1. Fitting orbit of Ceres (over-determi... ser- v...
2. T... distance to Cygni — Bessel
3. . . .

Little Used in the XIX Century

Nievergelt, “A tutorial history of least squares with applications to astronomy and geodesy,” in *Numerical Analysis: Historical Developments in the 20th Century*, edited by Brezinski and Wuytack, North Holland, 2001.

Recurring Work:

1. Ephemerides
2. Nautical Almanacs
3. Reducing observations to celestial coordinates

Grier, *When Computers Were Human*, Princeton, 2005.

Observations Needing Adjustment

case 1. Over-determined observations

Little Used in the XIX Century

$$\min_{x} \|b - Ax\|_2$$

solved by $A^t A x = c$ with $c = A^t b$

case 2. Under-determined observations

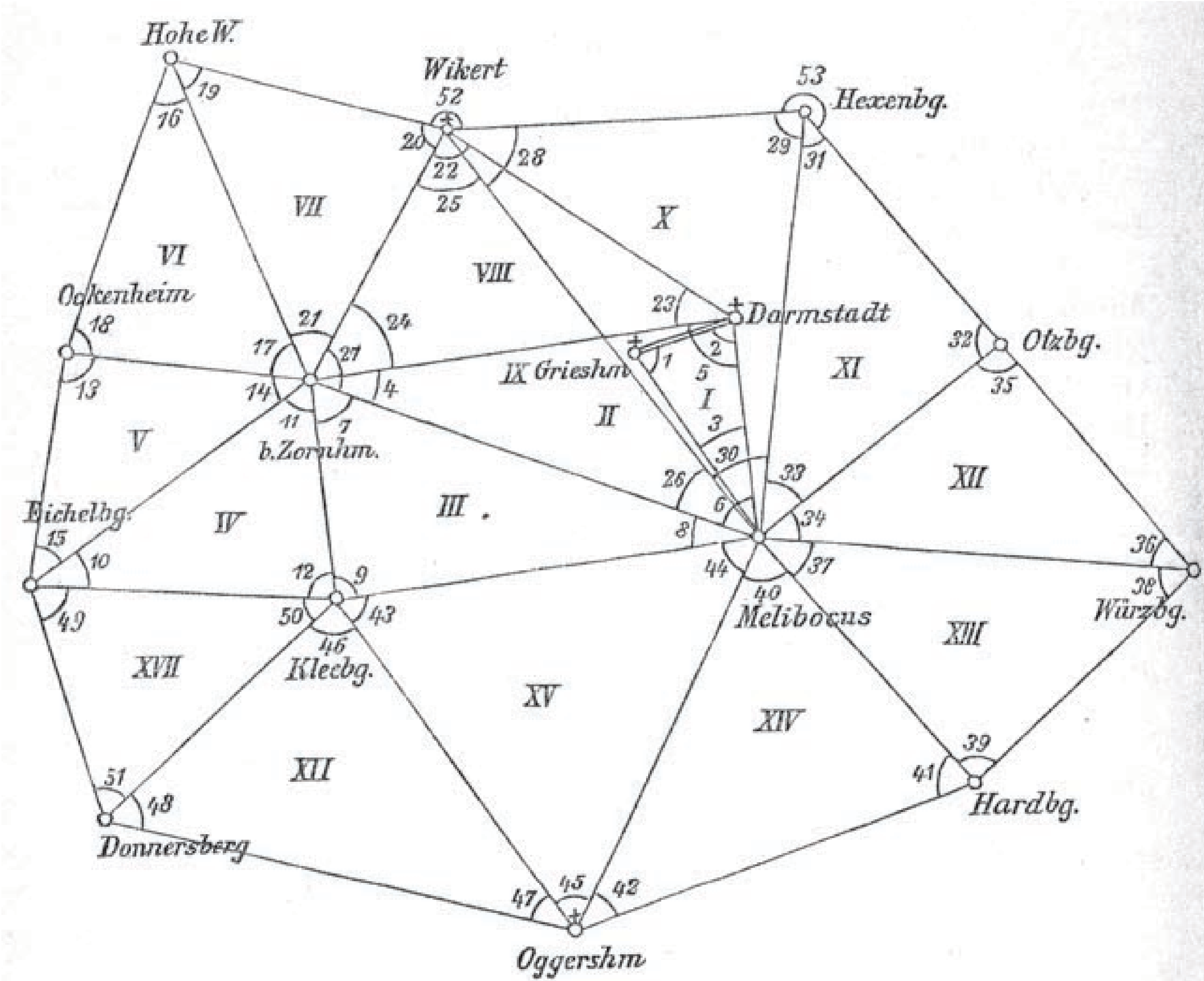
$$\min_{Ax = b} \|x\|_2$$

solved by $AA^t u = b$ where $x = A^t u$

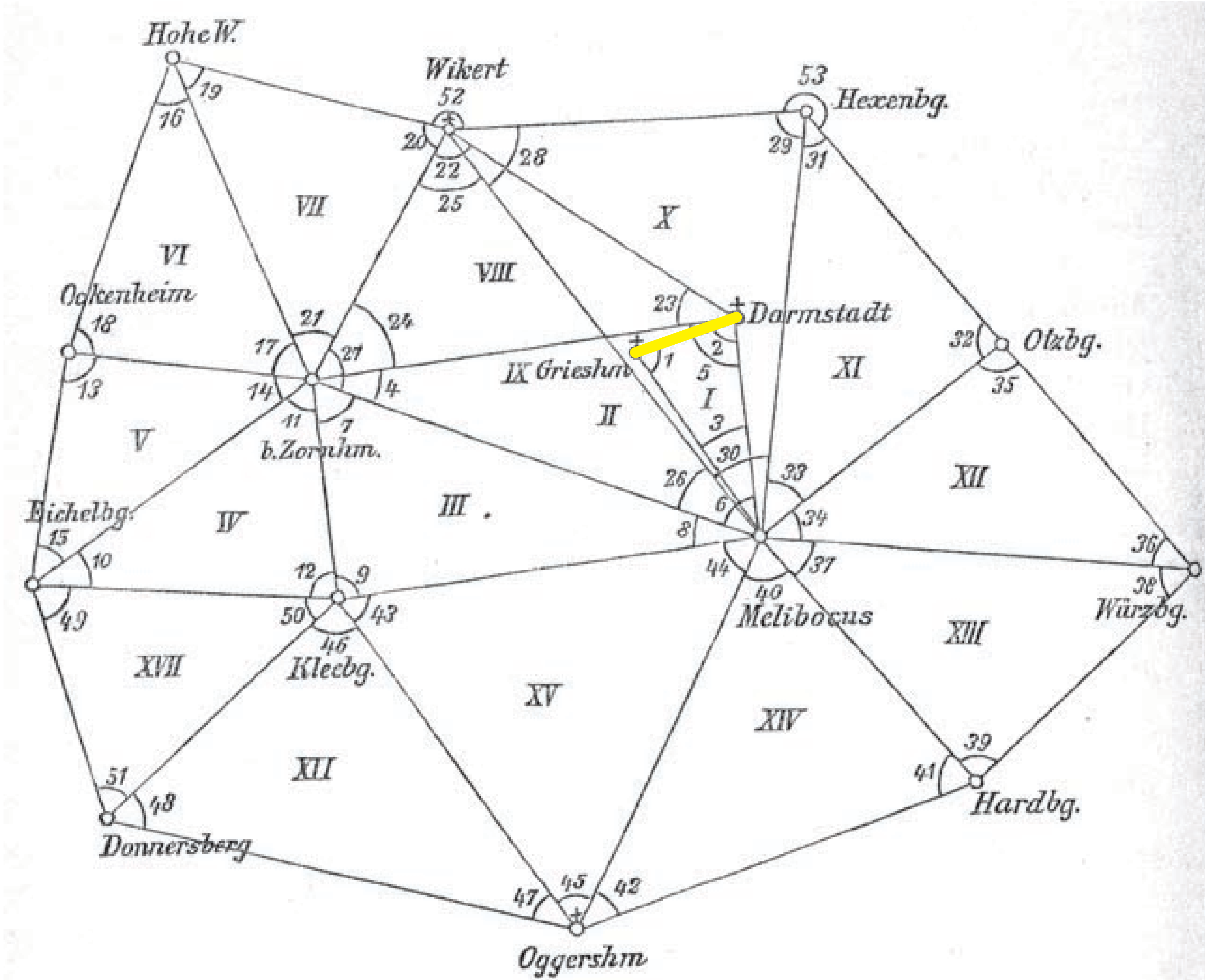
Both resulted in *normal equations*

Survey Adjustments

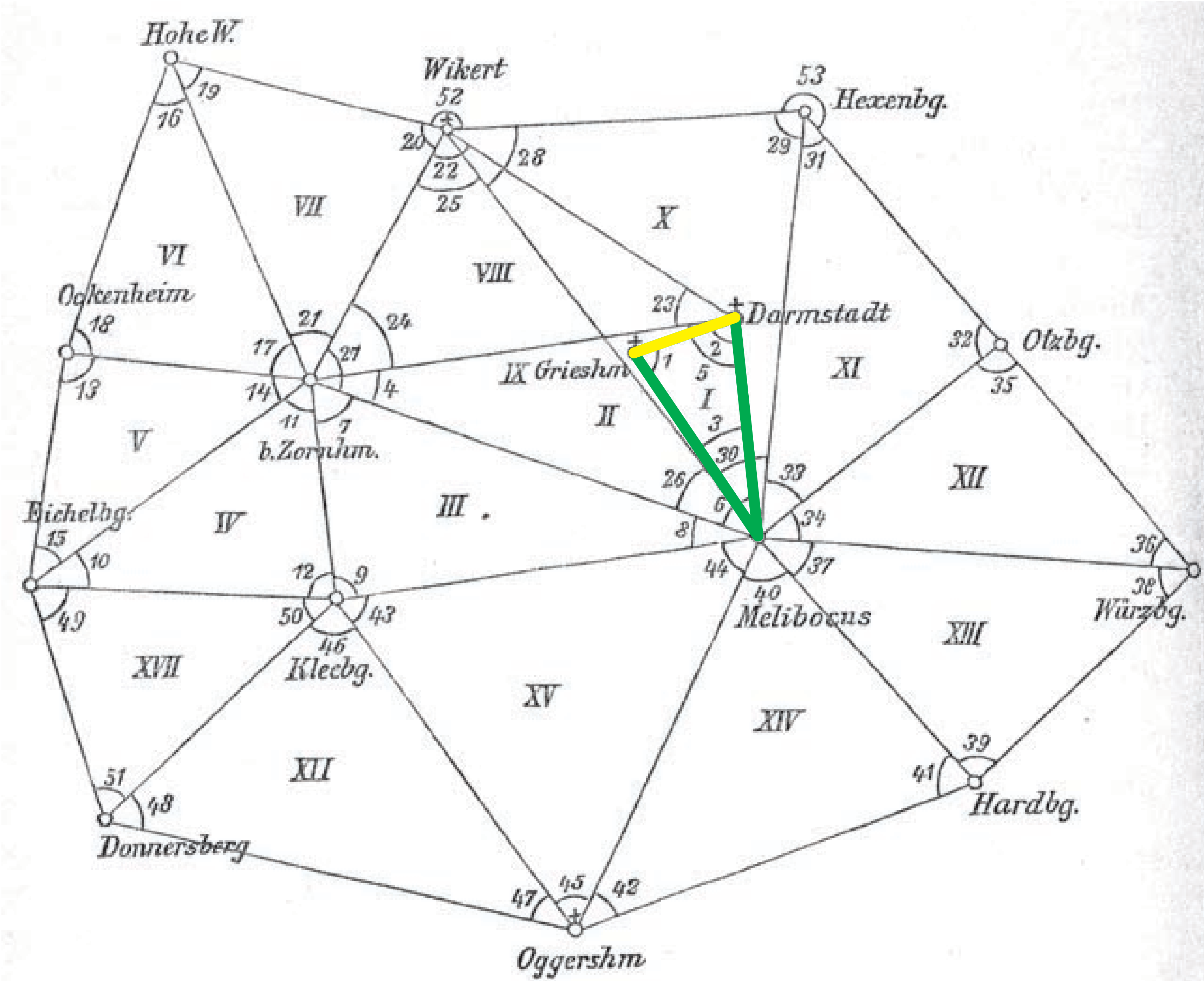
1880 Surveyor's Triangulation Net



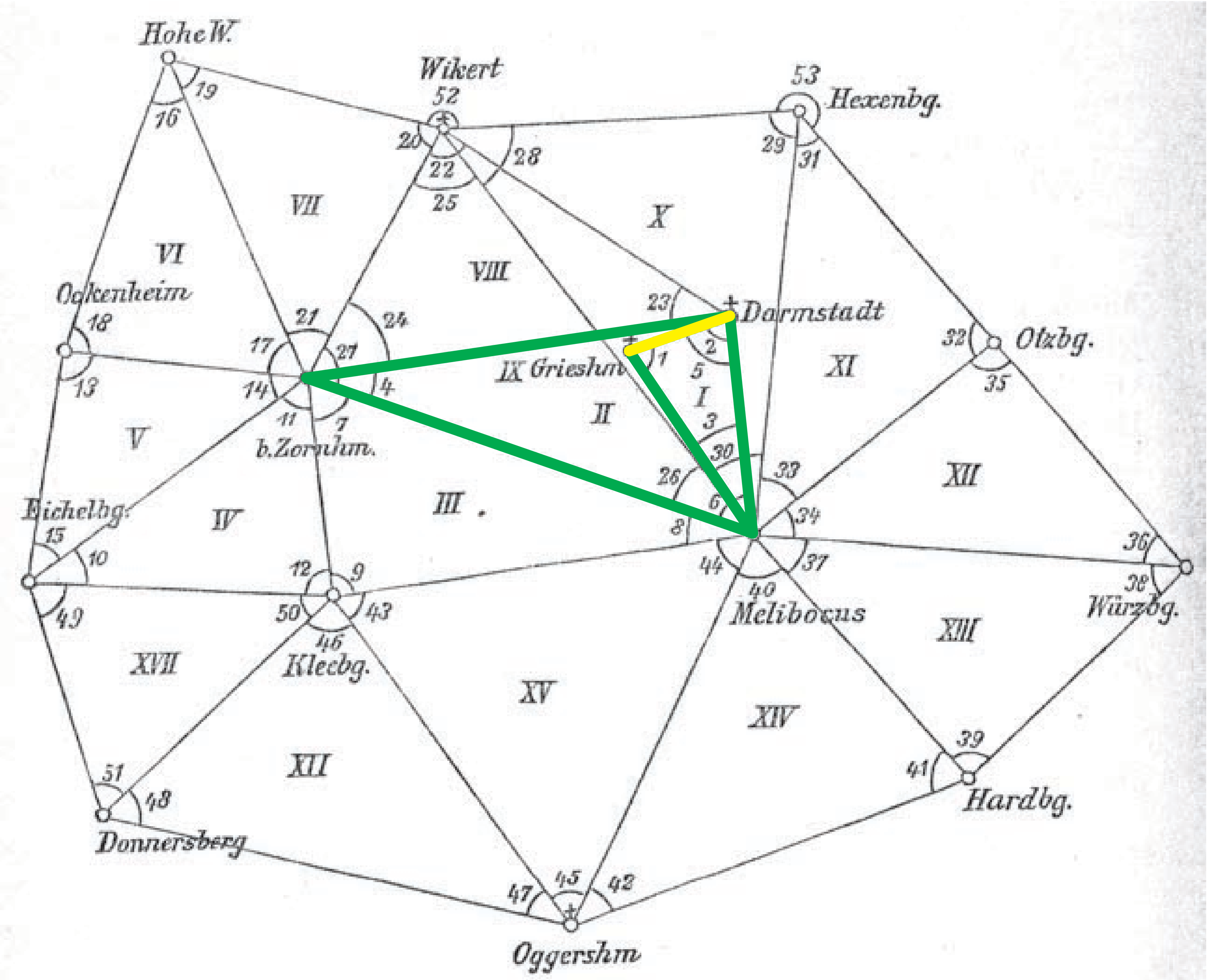
1880 Surveyor's Triangulation Net



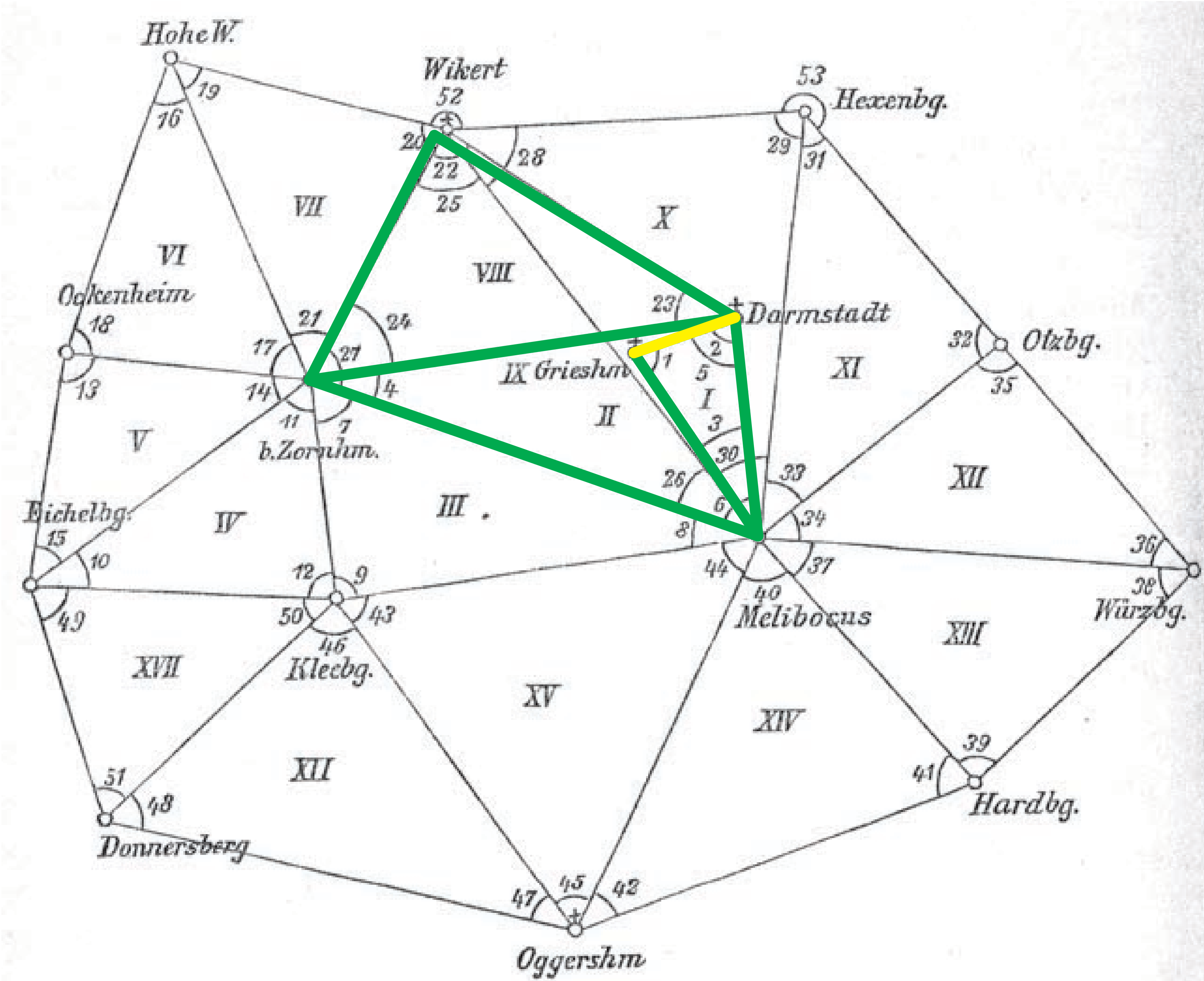
1880 Surveyor's Triangulation Net



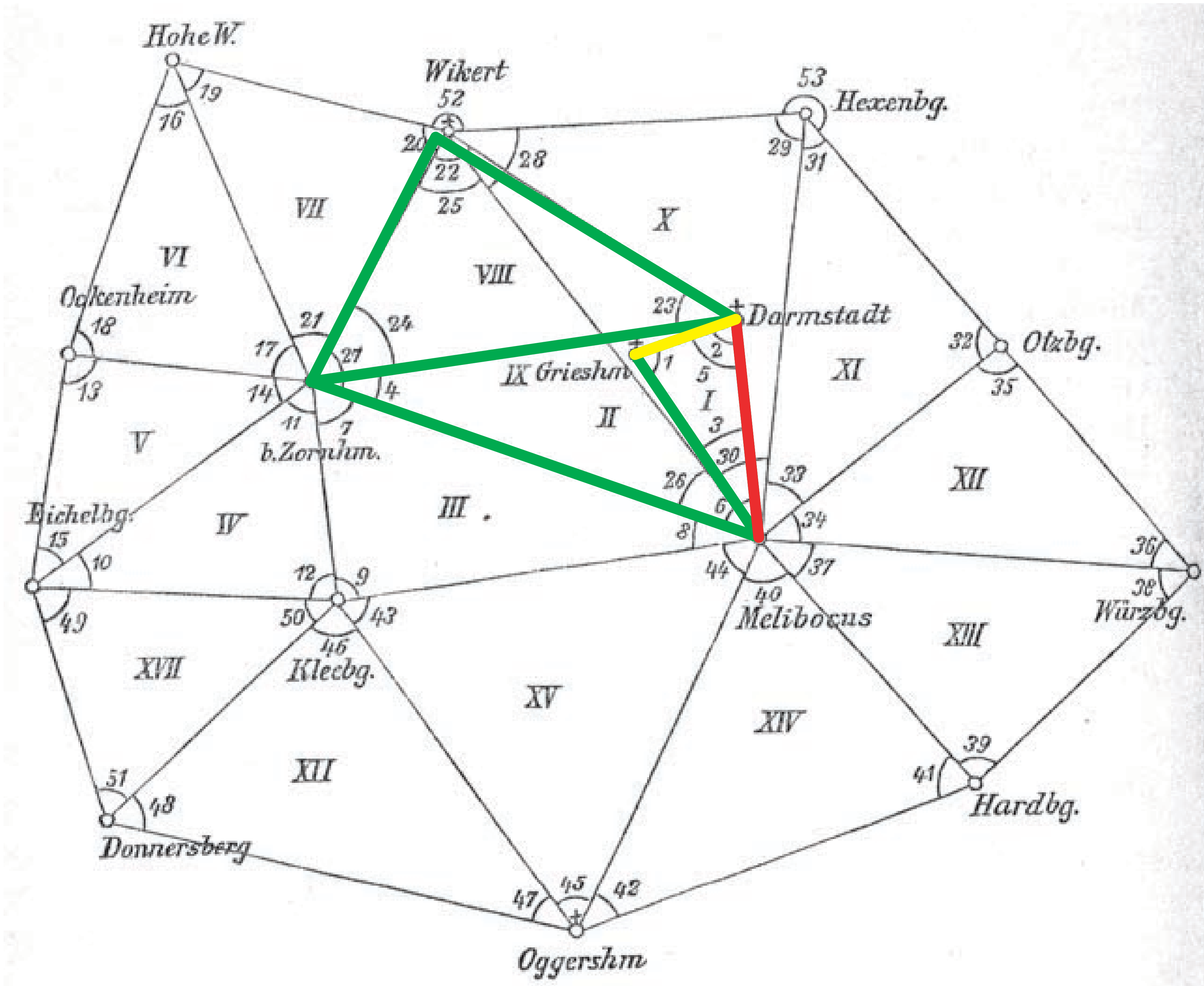
1880 Surveyor's Triangulation Net



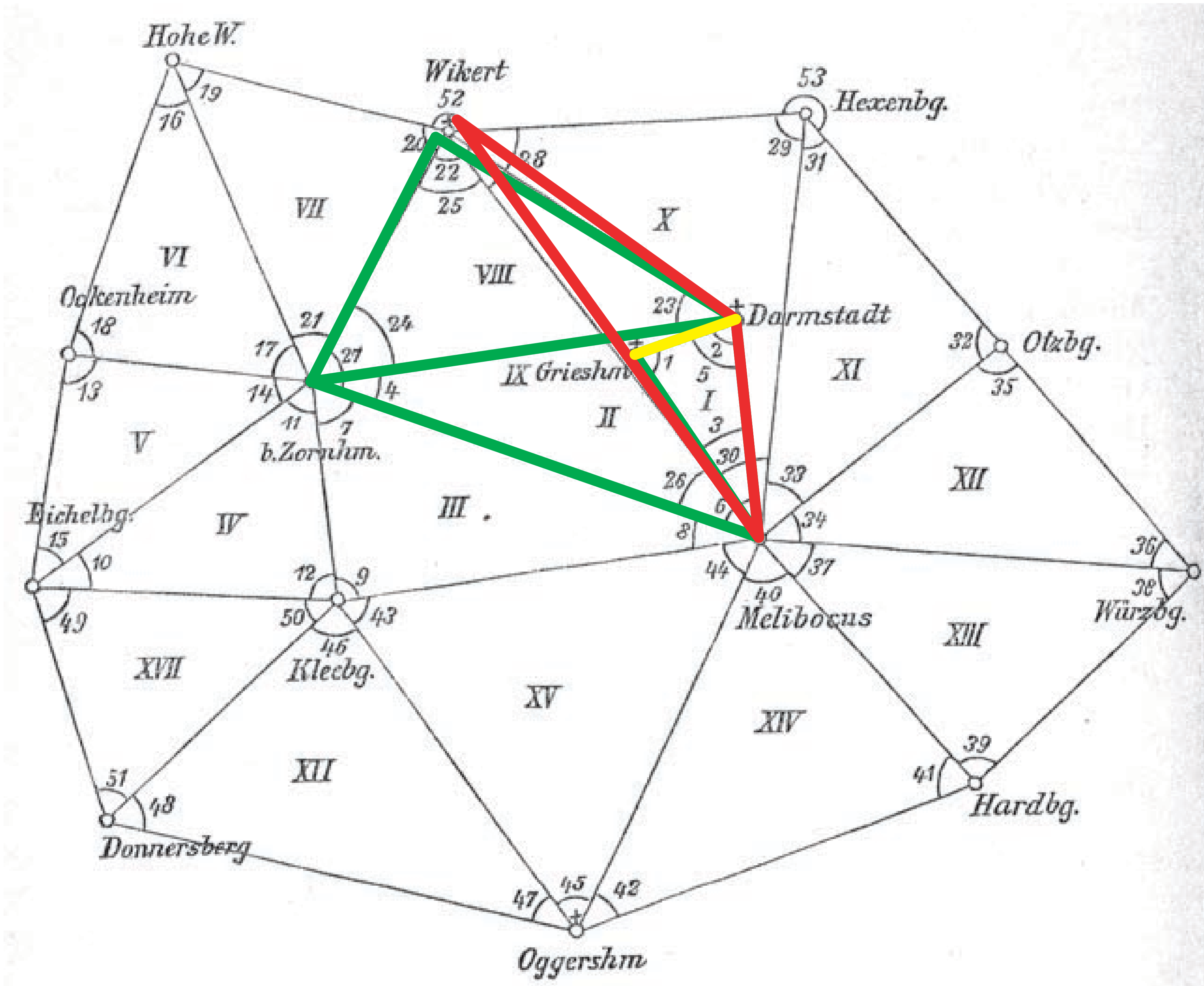
1880 Surveyor's Triangulation Net



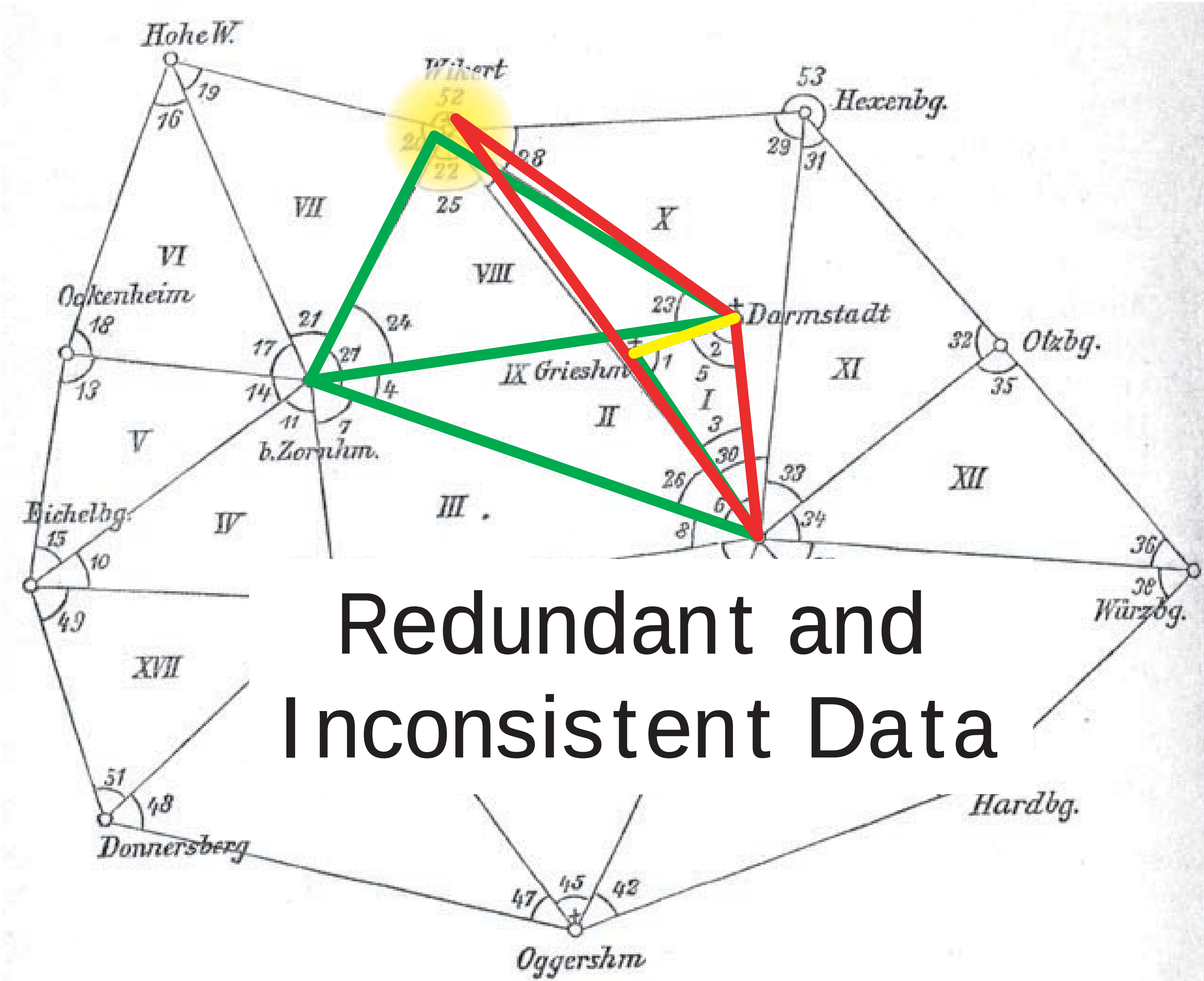
1880 Surveyor's Triangulation Net



1880 Surveyor's Triangulation Net



1880 Surveyor's Triangulation Net



Redundant and Inconsistent Data

Angle Adjustment Problem Addressed by

case 2. Under-determined observations

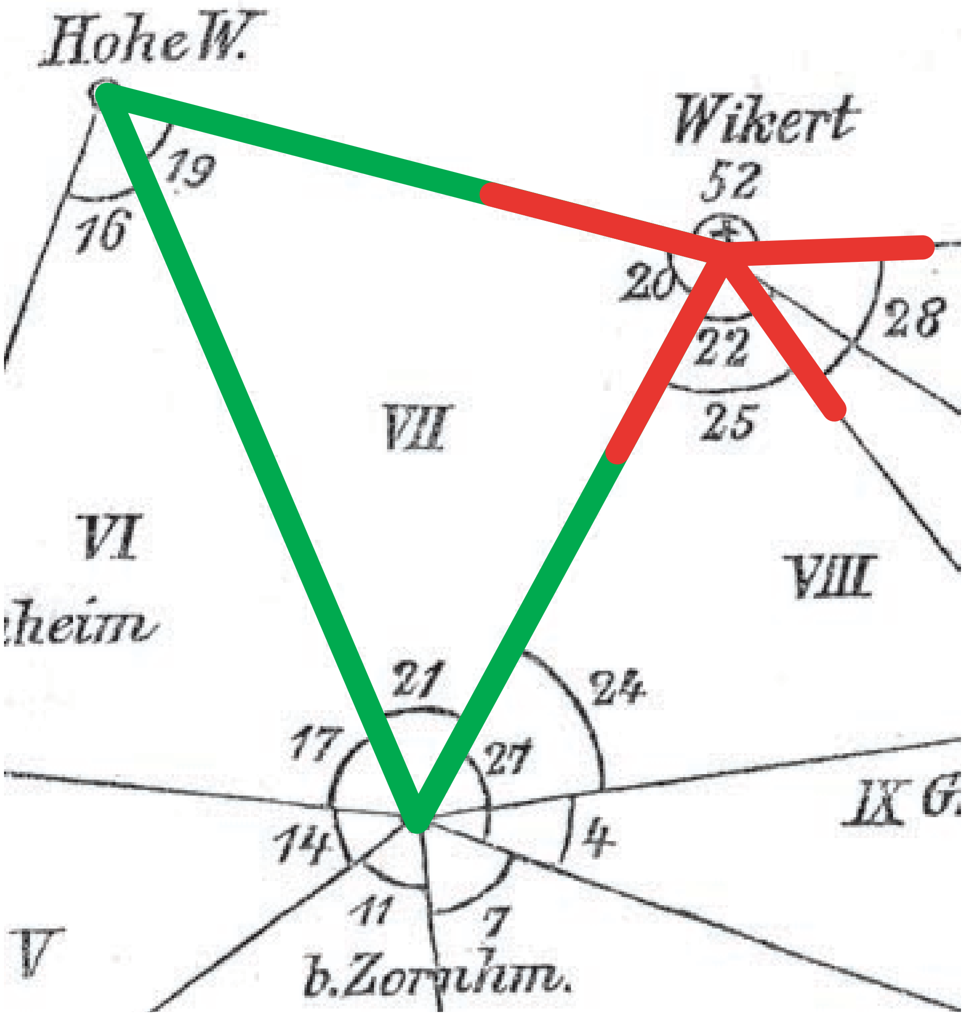
$$\min_{\boldsymbol{Ax} = \boldsymbol{b}} ||\boldsymbol{x}||_2$$

solved by $\boldsymbol{AA^t u} = \boldsymbol{b}$ where $\boldsymbol{x} = \boldsymbol{A^t u}$

- Unknowns $\boldsymbol{x}$ were adjustments to all (or most of) the measured angles
- Constraints $\boldsymbol{Ax} = \boldsymbol{b}$ called *conditions*

1. angle conditions (2 kinds)
2. side conditions

1. “Angle Equation” Conditions



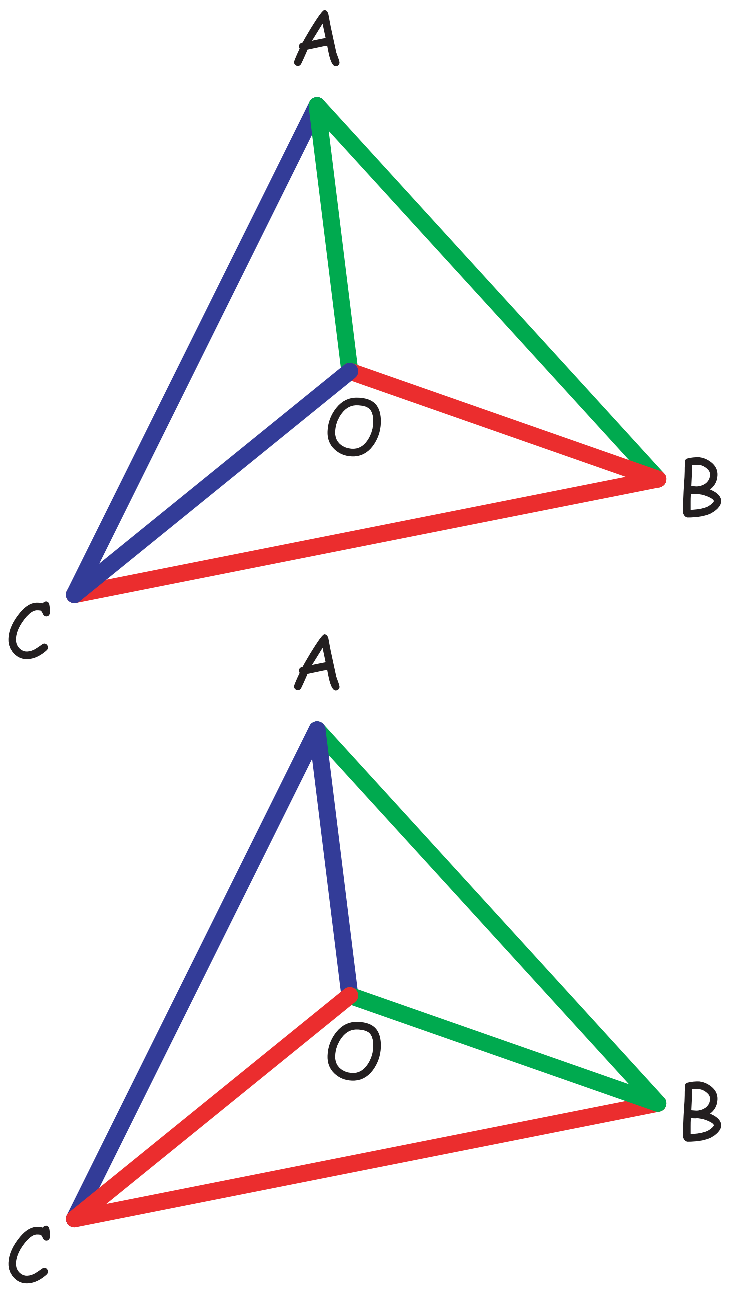
$\angle\overline{\mathbf{52}} = $ measured angle

$\epsilon_{52} = $ adjustment

$$\angle\overline{\mathbf{52}} + \epsilon_{52} + \angle\overline{\mathbf{28}} + \epsilon_{28} + \angle\overline{\mathbf{25}} + \epsilon_{25} + \angle\overline{\mathbf{20}} + \epsilon_{20} = \mathbf{360}$$

$$\angle\overline{\mathbf{19}} + \epsilon_{19} + \angle\overline{\mathbf{20}} + \epsilon_{20} + \angle\overline{\mathbf{21}} + \epsilon_{21} = \mathbf{180} + \text{correction}$$

2. “Side Equation” Condition (Sine Laws)

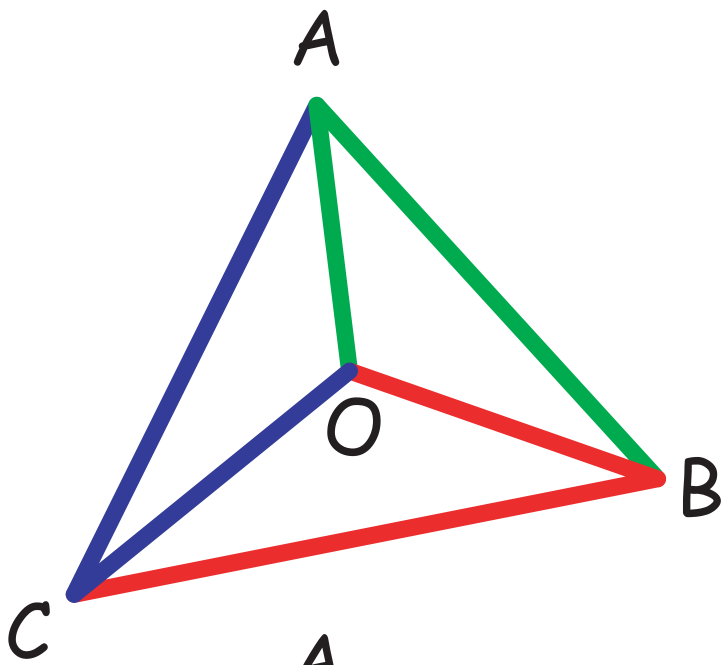


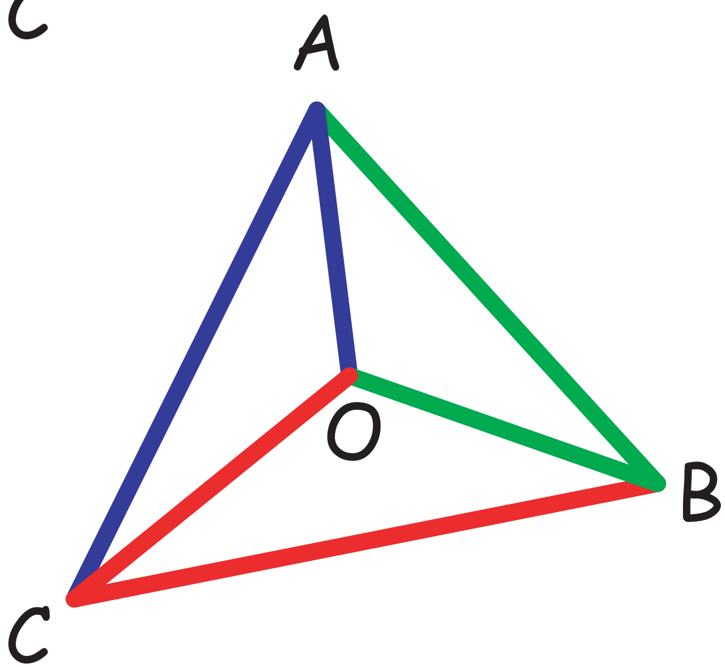
$$\sin(\overline{OAB} + \epsilon_1) \times \sin(\overline{OBC} + \epsilon_2) \times \sin(\overline{OCA} + \epsilon_3) = \sin(\overline{OBA} + \epsilon_4) \times \sin(\overline{OCB} + \epsilon_5) \times \sin(\overline{OAC} + \epsilon_6)$$

$\overline{OAB}$ = measured angle

ϵ_1 = adjustment

2. Side Equation Condition





$$\ln\sin\overline{OAB} + \epsilon_1 \cot\overline{OAB} +$$
$$\ln\sin\overline{OBC} + \epsilon_2 \cot\overline{OBC} +$$
$$\ln\sin\overline{OCA} + \epsilon_3 \cot\overline{OCA}$$
$$=^*$$
$$\ln\sin\overline{OBA} + \epsilon_4 \cot\overline{OBA} +$$
$$\ln\sin\overline{OCB} + \epsilon_5 \cot\overline{OCB} +$$
$$\ln\sin\overline{OAC} + \epsilon_6 \cot\overline{OAC}$$

* Equality is only to first order in the angle corrections.

Die Coefficienten der Bedingungsgleichungen

	a	b	c	d	e	f	g	h	i	k	l	m	n	o	p	q	r	s	t		w	x	y	z
1	0	0	0	0							1													
2	0	0	0	0							1													
3	0	0	0	0							1													
4	0	0	0	0	1	1						1												
5	cot 5	cot 5	0	0								1												
6	— cot 6	— cot 6	0	0								1												
7	0	0	— cot 7	cot 7	1								1											
8	cot 8	0	0	— cot 8			1						1											
9	— cot 9	0	cot 9	0				1					1											
10	— cot 10	0	0	cot 10										1										
11	0	0	0	— cot 11	1									1										
12	cot 12	0	0	0				1						1										
13	— cot 13	0	0	0											1									
14	0	0	0	0	1										1									
15	cot 15	0	0	0											1									
16	— cot 16	0	0	0												1								
17	0	0	0	0	1											1								
18	cot 18	0	0	0												1								
19	cot 19	0	0	0													1							
20	— cot 20	0	0	0					1								1							
21	0	0	0	0	1												1							
22	cot 22	cot 22	0	0														1						
23	— cot 23	— cot 23	0	0														1						
24	0	0	0	0	1	1												1						
25	0	— cot 25	— cot 25	0					1										1					
26	0	cot 26	0	0			1												1					
27	0	0	cot 27	0		—1													1					
28	0	0	cot 28	0					1											1				
29	0	0	— cot 29	0						1										1				
30	0	0	0	0			1													1				
31	0	0	cot 31	0						1											1			
32	0	0	— cot 32	0																	1			
33	0	0	0	0			1														1			
34	0	0	0	0			1															1		
35	0	0	cot 35	0																		1		
36	0	0	— cot 36	0																		1		

Gaussian Elimination

1823, 1828 Gauss

J. C. F. Gauss

1777 – 1855

Theoria combinationes observationum erroribus minimis obnoxiae: Pars prior and *Supplementum*, Göttingen, 1823 and 1828.

case 2.

Under-determined observations

$$\min_{Ax=b} \|x\|_2$$

solved by $AA^t u = b$

where $x = A^t u$.

1823, 1828 Gauss

J. C. F. Gauss

1777 – 1855

Theoria combinationes observationum erroribus minimis obnoxiae: Pars prior and *Supplementum*, Göttingen, 1823 and 1828.

Normal equations:

$$
\begin{aligned}
0 &= [aa]x + [ab]y + [ac]z + \\
0 &= [ab]x + [bb]y + [bc]z + \\
0 &= [ac]x + [bc]y + [cc]z + \\
0 &= [ad]x + [bd]y + [cd]z + \\
&\text{etc.}
\end{aligned}
$$

Stewart, “Gauss, Statistics, and Gaussian Elimination,” in *Computing Science and Statistics: Computationally Intensive Statistical Methods*, edited by Sall and Lehman, Fairfax Station, 1994

1823, 1828 Gaussian Elimination

J. C. F. Gauss

1777 – 1855

Theoria combinationes observationum erroribus minimis obnoxiae: Pars prior and *Supplementum*, Göttingen, 1823 and 1828.

$$[bb,1] = [bb] - \frac{[ab]^2}{[aa]}$$

$$[bc,1] = [bc] - \frac{[ab][ac]}{[aa]}$$

etc.

$$[cc,2] = [cc] - \frac{[ac]^2}{[aa]} - \frac{[bc,1]^2}{[bb,1]}$$

$$[cd,2] = [cd] - \frac{[ac][ad]}{[aa]} - \ldots$$

etc.

Dana P. Bartlett

General Principles of the Method of Least Squares with Applications, third edition, Boston, 1915.

Normal equations:

$$\begin{aligned}
0 &= [aa]z_1 + [ab]z_2 + [ac]z_3 \\
0 &= [ab]z_1 + [bb]z_2 + [bc]z_3 \\
0 &= [ac]z_1 + [bc]z_2 + [cc]z_3 \\
0 &= [ad]z_1 + [bd]z_2 + [cd]z_3
\end{aligned}$$

etc.

1915 Dana Bartlett

Dana P. Bartlett

General Principles of the Method of Least Squares with Applications, third edition, Boston, 1915.

1^{st} reduced normal equations:

$$[bb,1]z_2 + [bc,1]z_3 + [bd,1]$$

$$[bc,1]z_2 + [cc,1]z_3 + [cd,1]$$

$$[bd,1]z_2 + [cd,1]z_3 + [dd,1]$$

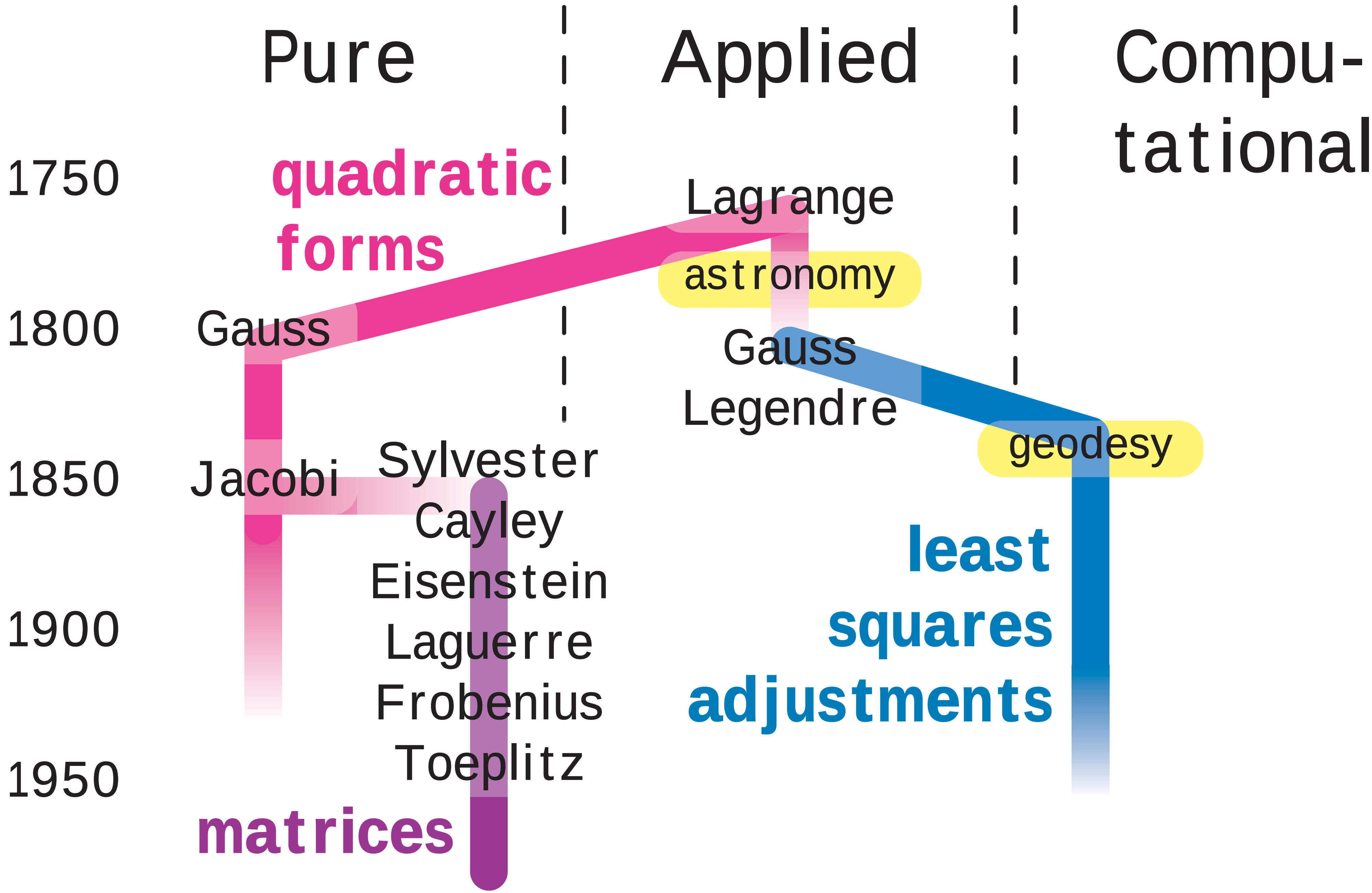
$$\cdots$$

where

$$[bb,1] = [bb] - \frac{[ab][ab]}{[aa]}$$

$$[bc,1] = [bc] - \frac{[ab][ac]}{[aa]}$$

Linear Algebra Timeline



Gauss solved the normal equations

- Gauss invented a notation to organize elimination that was used for 100 years.
- During most of its existence, Gaussian elimination was used to solve normal equations.

Pertinent to the Inversion Paper

- Von Neumann and Goldstine studied the rounding errors of

$$A^{-1} = A^t(AA^t)^{-1}$$

  This essentially was the main use for Gaussian elimination for 150 years.

- Their results stimulated a search for new algorithms (QR, SVD) that eventually replaced the normal equations in 1965.

Geodesy

Surveying Continents

Southeast Alaska
1929

1. Surveyors
   - gather baseline and angle data
2. Computers (original usage)
   (a) form underdetermined angle adjustment eqns
   (b) form and solve normal equations
3. Cartographers
   - use adjusted triangulations to draw maps

Surveying Continents

US center for numerical linear algebra through WWI

Old C&GS Office with third floor computing rooms

Washington

1. Surveyors
   - gather baseline and angle data
2. Computers (original usage)
   (a) form underdetermined angle adjustment eqns
   (b) form and solve normal equations
3. Cartographers
   - use adjusted triangu-lations to draw maps

U. S. Coast and Geodetic Survey (C&GS)

Oldest American Scientific Agency

Benjamin Peirce

3rd Superintendent

(served 1867 – 74)

Charles Schott

Assistant

in charge of the

Computing Division

(served 1848 – 99)

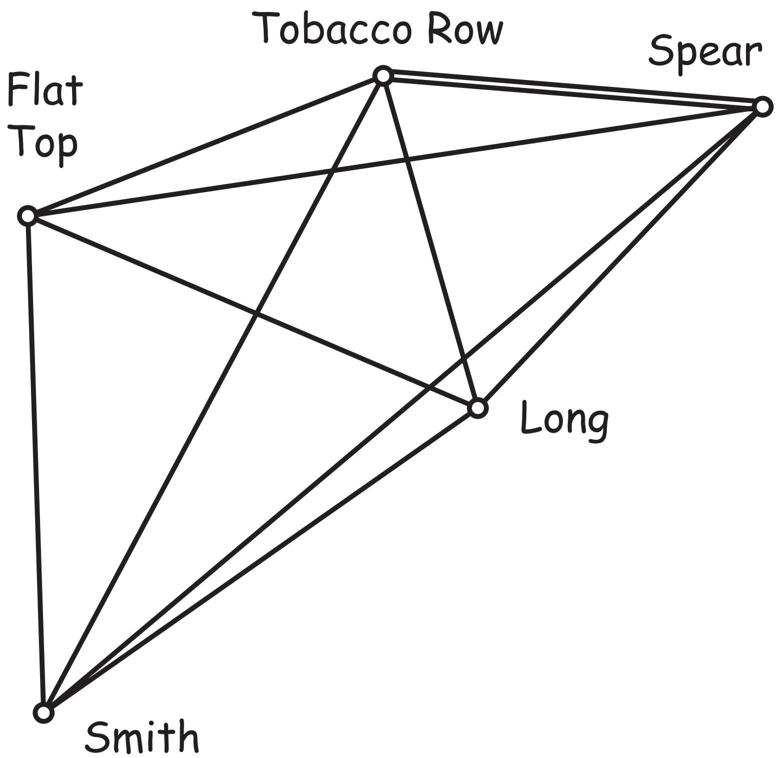
Myrrick Doolittle

1830 – 1911

Computer

(served 1873 – ??)

The Numerical Linear Algebra



Doolittle's Example

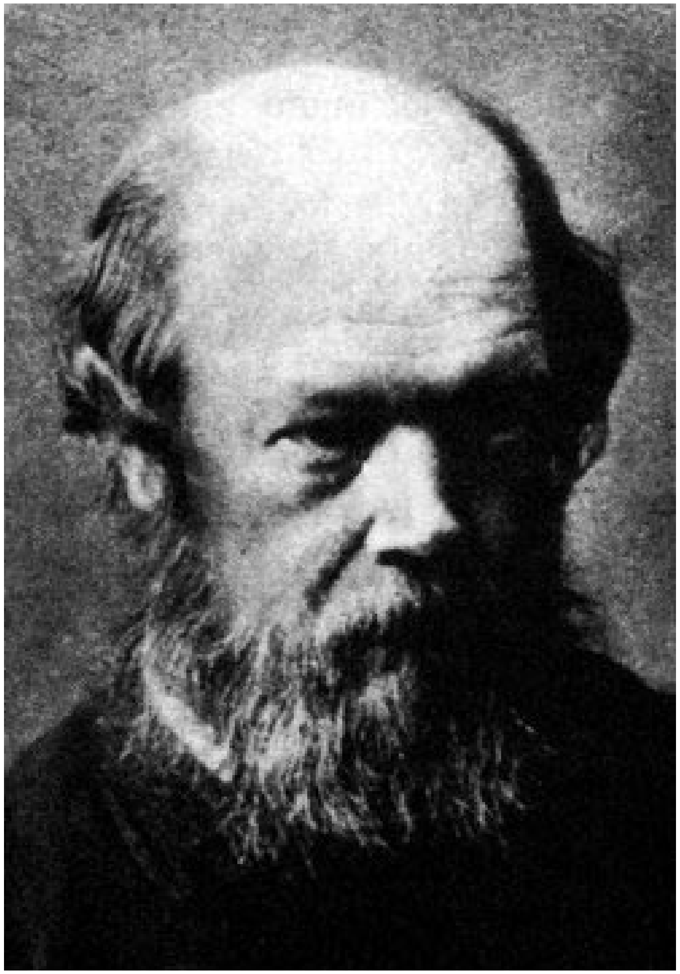
Doolittle, "Method employed in the solution of normal equations and the adjustment of a triangularization," in *Report of the Superintendent of the U. S. Coast and Geodetic Survey (showing the progress of the work during the fiscal year ending with June, 1878)*, Government Printing Office, 1881.

Selection of Conditions:

- $\mathbf{2}$ unknowns for each *station* off the baseline
- What is the rank of the matrix of conditions?
- $\boldsymbol{n}$ as big as $\mathbf{41}$ or $\mathbf{52}$!

Solution of Normal Equations:

- 3-digit hand arithmetic!
- Crelle's 3-digit tables were used for $\times$ and $\div$



Peter Guthrie Tait

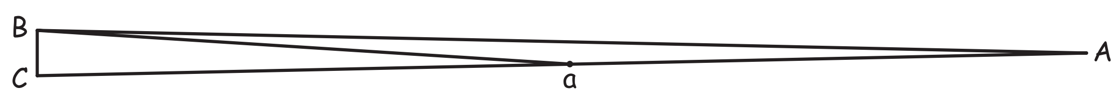
1831 – 1901

"Cosmical Astronomy," Good Words, 19–23 (1875).

"The danger of selecting such condition equations that the solution will be somewhat unstable is difficult to avoid."

Wright and Hayford, "The Adjustment of Observations by the Method of Least Squares with Applications to Geodetic Work," Van Nostrand, New York, 1906.

"The difficulty of determining the sun's distance lies in the measurement of the angles, and in what is called the *ill-conditioned* form of triangle."



Pertinent to the Inversion Paper

- Von Neumann and Goldstine’s error bounds were considered unrealistically large when tested on small problems solved by *post*-WWII technology (10-digit mechanical calculators) . . .

- . . . but the concerns about Gaussian elimination were based on experience from *pre*-WWII calculations (3-digit paper and pencil) for which the pessimistic bounds probably were realistic.

Another Story: Regression and Correlation

Karl Pearson
1857 – 1936

Statistics replaced geodesy as the main use for linear algebra after WWI

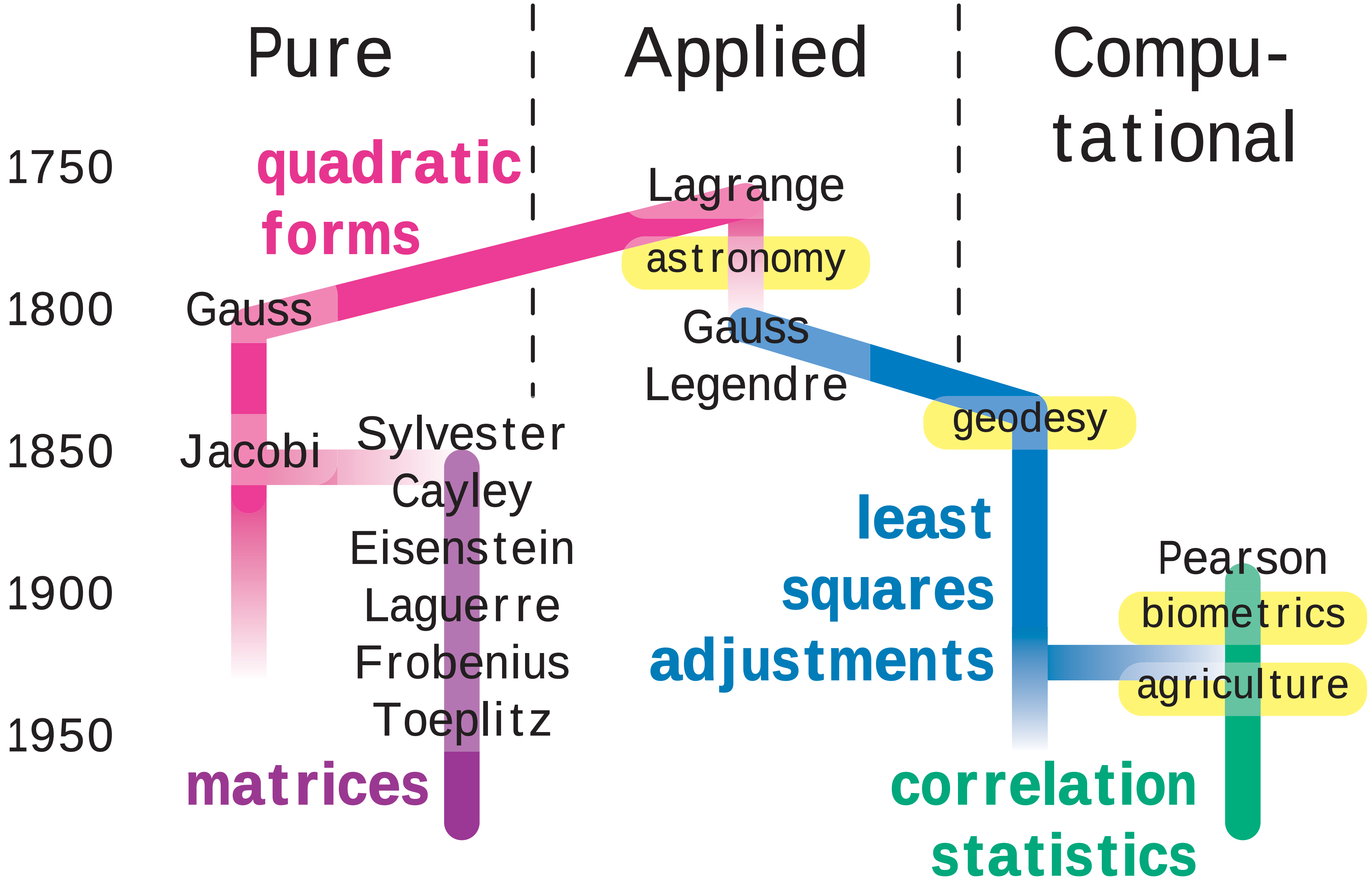
- continental surveys were complete
- data analysis was important for agriculture and government

Karl Pearson led in the computational developments and published *Tracts for Computers*, essentially the first numerical analysis textbooks

Grier, *When Computers Were Human*, Princeton, 2005.

Question: actuaries also used computing; did they use linear algebra?

Linear Algebra Timeline



Pertinent to the Inversion Paper

- Von Neumann and Goldstine deferred to Hotelling in matters of numerical linear algebra because, as a leader of the American statistical community, he represented a profession with considerable computing expertise.

Numerical Linear Algebra Algorithms Correspond to Matrix Factorizations

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$$

Who invented the paradigm?

Stewart, "The Decompositional Approach to Matrix Computation," Computing in Science and Engineering, 2(1):50–59 (2000).

1823, 1828 Gaussian Elimination

J. C. F. Gauss

1777 – 1855

Theoria combinationes observationum erroribus minimis obnoxiae: Pars prior and *Supplementum*, Göttingen, 1823 and 1828.

$$[bb,1] = [bb] - \frac{[ab]^2}{[aa]}$$

$$[bc,1] = [bc] - \frac{[ab][ac]}{[aa]}$$

$$\text{etc.}$$

$$[cc,2] = [cc] - \frac{[ac]^2}{[aa]} - \frac{[bc,1]^2}{[bb,1]}$$

$$[cd,2] = [cd] - \frac{[ac][ad]}{[aa]} - \ldots$$

$$\text{etc.}$$

Before matrices; no paradigm.

1924 Cholesky Factorization

Andre-Louis Cholesky

1875 – 1918

Benoit, “Note sur une méthode . . . ,” *Bulletin géodésique*, 67–77 (1924).

$$\min_{Ax=b} \|x\|_2$$

$AA^t u = b$ where $x = A^t u$.

Cholesky's insight was that many matrices have the same normal equations. He chose one that was easy to solve.

$$LL^t = AA^t$$

Real breakthrough was the inner product formulas for L.

No matrices in the paper, hence no paradigm.

1907 “Toeplitz” Factorization?

Toeplitz 1881 – 1940
Jacobi 1804 – 1851

Toeplitz, “Die Jacobische Transformation der quadratischen Formen von unendlichvielen Veränderlichen,” *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, 101–109 (1907).

Restated Jacobi’s result “known to Lagrange and Gauss” using “matrix calculus of Frobenius.”

A quadratic form

$$S = \sum_{i,k=1}^{n} \alpha_{i,k} x_i x_k$$

could be factored as

$$S^{-1} = U'U$$

or $S = U^{-1} U'^{-1}$

but U is given by determinants

1938 “Banachiewicz” Factorization?

Tadeusz
Banachiewicz
1882 – 1954

Astronomer and polymath who calculated first orbit of Pluto.

Expressed calculations using “Cracovian” matrix multiplication. Had a Cholesky-like method.

Inspired Jensen to do the same for Cayleyan matrix multiplication.

Many papers in Academy of Science of Poland were widely cited.

1938 Dwyer Square-Root Factorization?

Paul Dwyer

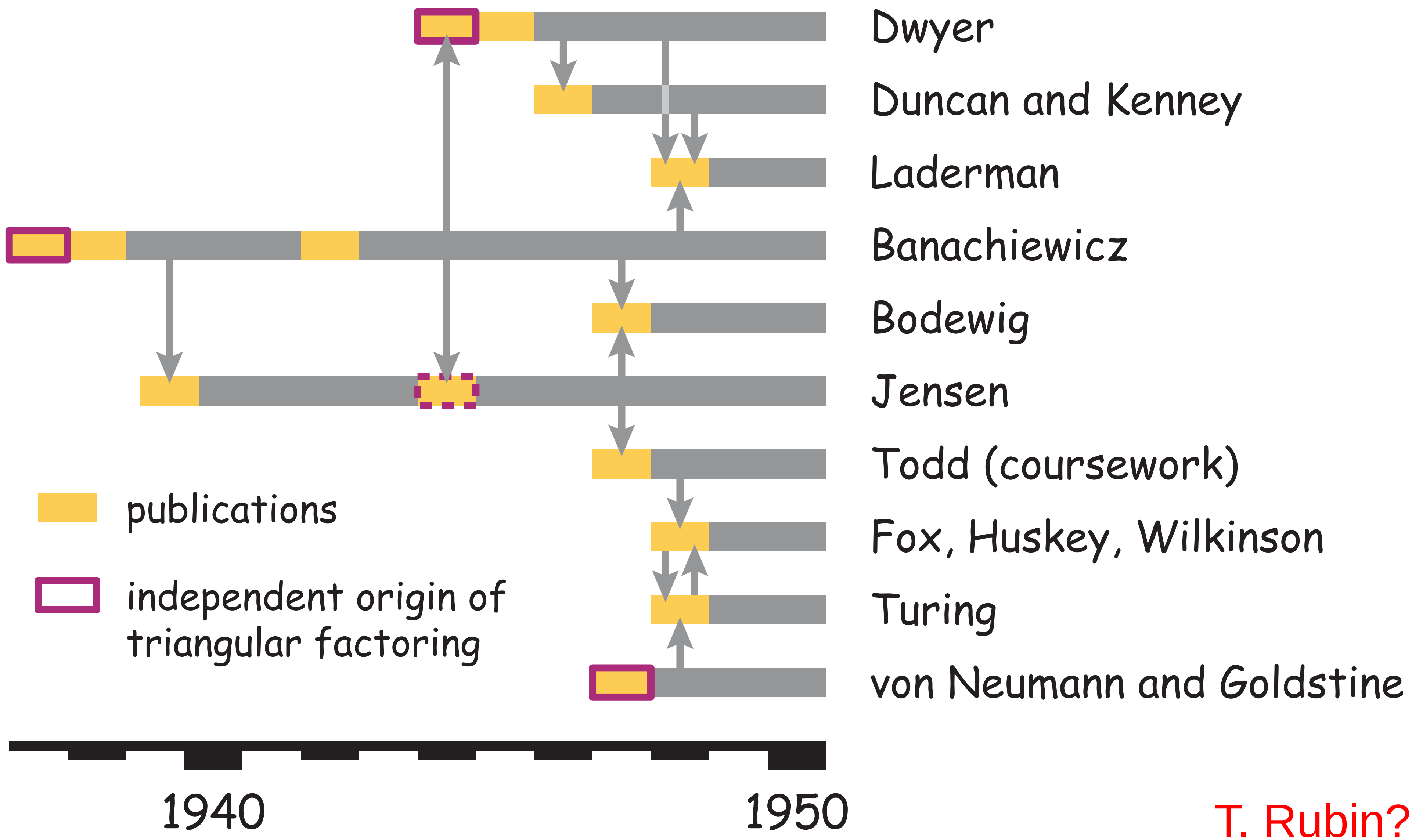
Dwyer, "A matrix presentation of least squares and correlation theory with matrix justification of improved methods of solution," *The Annals of Mathematical Statistics*, 5:82–89 (1944).

Michigan statistician.

Developed matrix formulation of least squares, normal equations, and symmetric factorization.

Primary source in the United States for what was later named "Cholesky's algorithm."

Citation Patterns Among Earliest Authors



T. Rubin?

- All other authors were interested in normal equations.
- Von Neumann and Goldstine were the first to describe factoring of nonsymmetric matrices. They originated the $\boldsymbol{LDU}$ form.

Numerical Linear Algebra Before the Last Paradigm Shift*

Joseph F. Grcar
Lawrence Berkeley Laboratory
jfgrcar@lbl.gov

October 5, 2005
Berkeley

*Thomas Kuhn, *The Structure of Scientific Revolutions*, The University of Chicago Press, Chicago, 1970.