John von Neumann's Analysis of Gaussian Elimination

and the Invention of Modern Computing

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

http://seesar.lbl.gov/ccse/people/grcar/Talks/talks.html

Outline

History of Gaussian Elimination

Invention of Electronic Computers

The Matrix Inversion Paper

Conclusion

Errors of Observation

Adrien-Marie Legendre

1752 – 1833

Johann Carl Friedrich Gauss

1777 – 1855

Pierre-Simon Laplace

1749 – 1827

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

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

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

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

case 2. Under-determined observations

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

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

Both solved by *normal equations*

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 $c = A^t b$

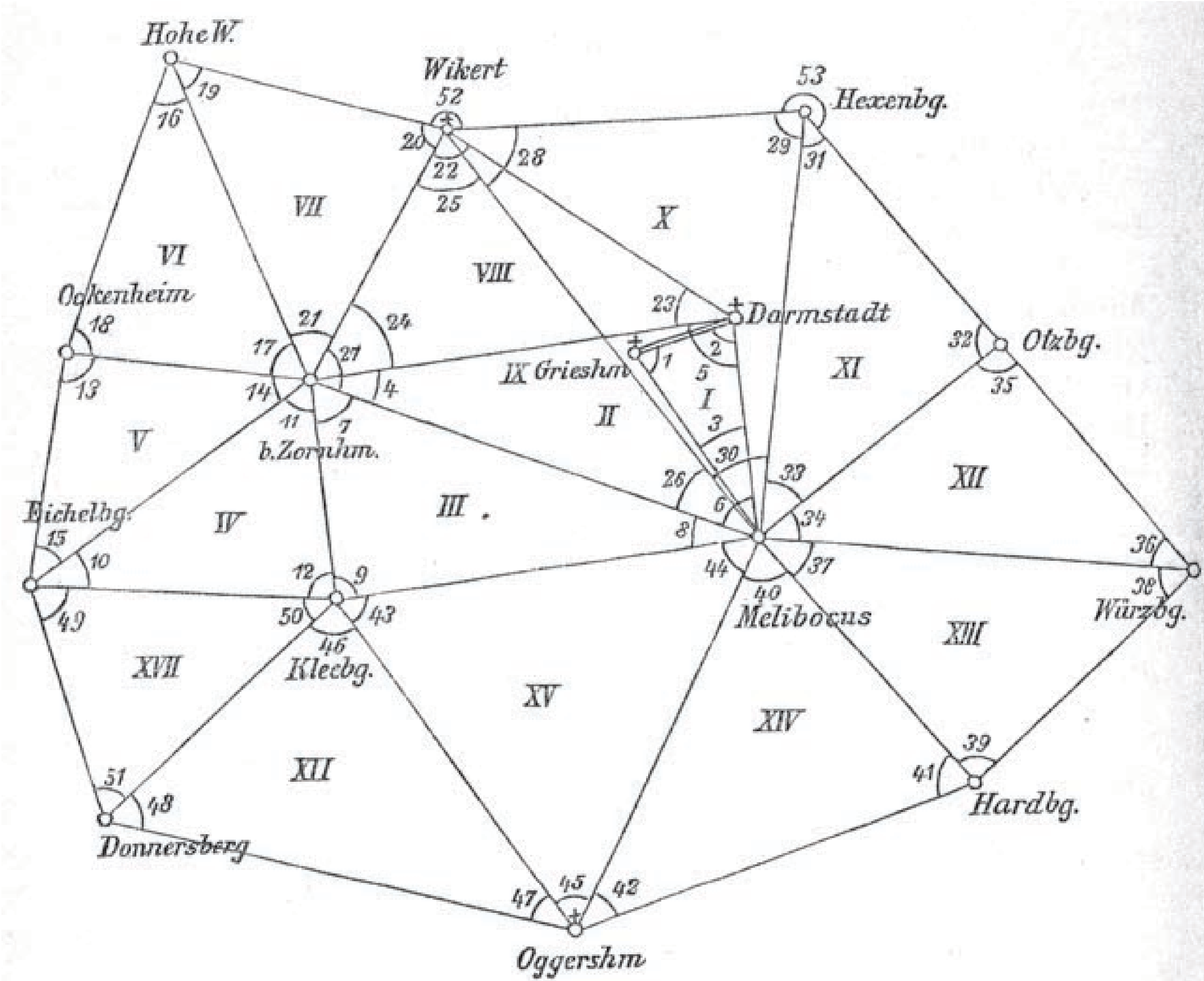
case 2. Under-determined observations

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

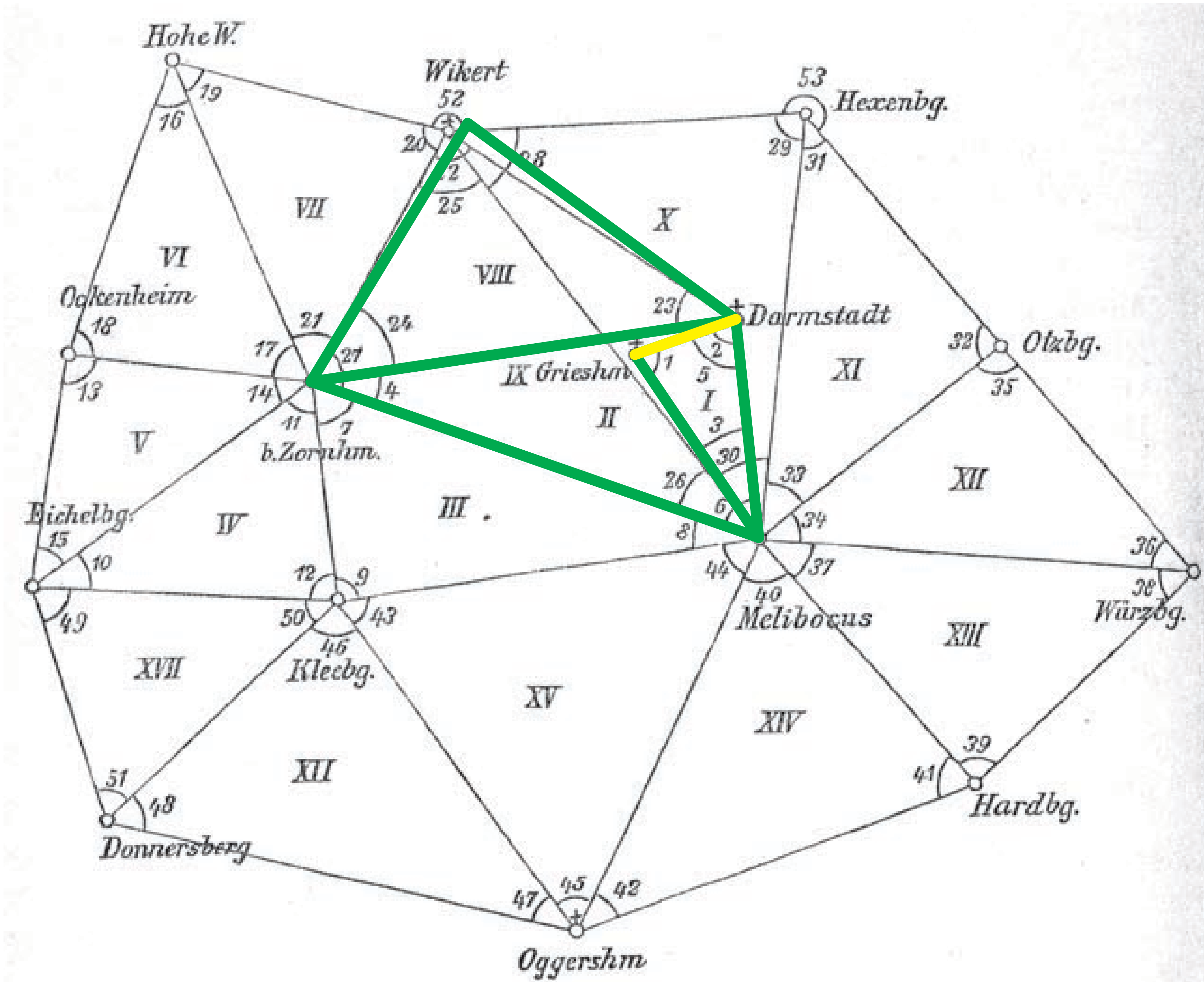
solved by $A A^t u = b$ where $x = A^t u$

Both solved by *normal equations*

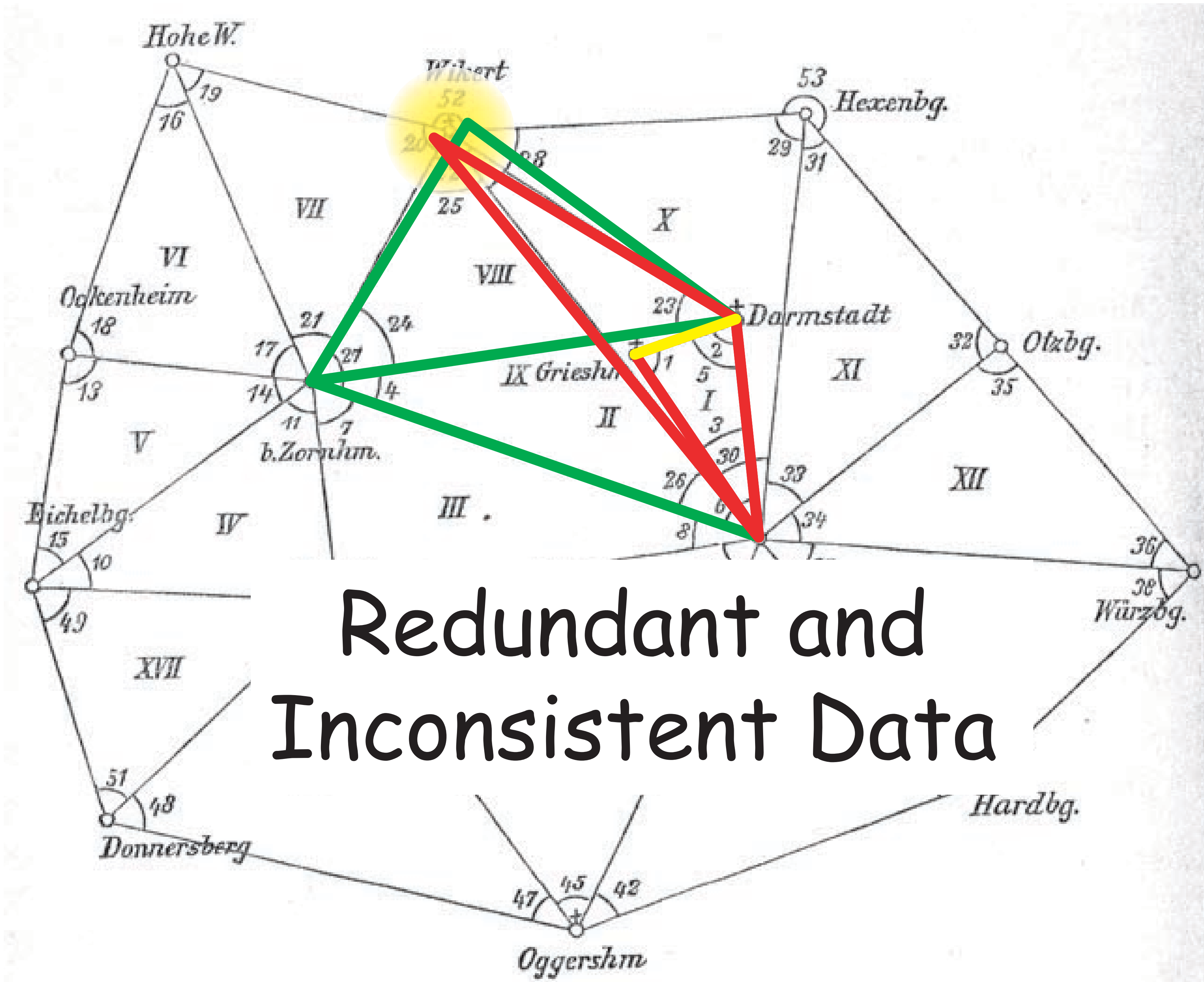
1880 Surveyor's Triangulation Net



1880 Surveyor's Triangulation Net



1880 Surveyor's Triangulation Net



Redundant and Inconsistent Data

Geodetic Angle Adjustment

- Unknown adjustments, $\boldsymbol{x}$, to the angles
- case 2. Under-determined observations

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

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

- Constraints, $\boldsymbol{Ax} = \boldsymbol{b}$, called *conditions*

1. angle conditions
2. side conditions

Die Coefficienten der Bedingungsgleichungen

	a	*b*	*c*	*d.*	*e*	*f*
1	0	0	0	0		
2	0	0	0	0		
3	0	0	0	0		
4	0	0	0	0	1	1
5	$\cot 5$	$\cot 5$	0	0		
6	$-\cot 6$	$-\cot 6$	0	0		
7	0	0	$-\cot 7$	$\cot 7$	1	
8	$\cot 8$	0	0	$-\cot 8$		
9	$-\cot 9$	0	$\cot 9$	0		
10	$-\cot 10$	0	0	$\cot 10$		
11	0	0	0	$-\cot 11$	1	
12	$\cot 12$	0	0	0		
13	$-\cot 13$	0	0	0		
14	0	0	0	0	1	
15	$\cot 15$	0	0	0		
16	$-\cot 16$	0	0	0		
17	0	0	0	0	1	
18	$\cot 18$	0	0	0		
19	$\cot 19$	0	0	0		
20	$-\cot 20$	0	0	0		
21	0	0	0	0	1	
22	$\cot 22$	$\cot 22$	0	0		
23	$-\cot 23$	$-\cot 23$	0	0		
24	0	0	0	0	1	1
25	0	$-\cot 25$	$-\cot 25$	0		
26	0	$\cot 26$	0	0		
27	0	0	$\cot 27$	0		-1
28	0	0	$\cot 28$	0		
29	0	0	$-\cot 29$	0		
30	0	0	0	0		
31	0	0	$\cot 31$	0		
32	0	0	$-\cot 32$	0		
33	0	0	0	0		
34	0	0	0	0		
35	0	0	$\cot 35$	0		
36	0	0	$-\cot 36$	0		

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

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:

$$0 = [aa]x + [ab]y + [ac]z + $$
$$0 = [ab]x + [bb]y + [bc]z + $$
$$0 = [ac]x + [bc]y + [cc]z + $$
$$0 = [ad]x + [bd]y + [cd]z + $$

etc.

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.

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

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

etc.

90 Years Later — MIT Prof. Dana Bartlett

Dana P. Bartlett

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

1$^{\text{st}}$ reduced normal equations:

$$\begin{aligned}
&[bb,1]z_2 + [bc,1]z_3 + [bd,1] \\
&[bc,1]z_2 + [cc,1]z_3 + [cd,1] \\
&[bd,1]z_2 + [cd,1]z_3 + [dd,1] \\
&\qquad\cdots
\end{aligned}$$

where

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

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

Gaussian Elimination Summary

Liu Hui
≈ 260

al-Khwarizmi
≈ 800

Gauss
≈ 1820

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

  1. described <u>without matrices</u>, and
  2. used to solve <u>normal equations</u>.

19 – 20th Century Continental Surveys

Surveyor

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 triangu-lations to draw maps

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 triangulations 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 Super.

in charge of the

Computing Division

(served 1848 – 99)

Myrrick Doolittle

1830 – 1911

Computer

(served 1873 – ??)

After WWI — Statistics

Karl Pearson
1857 – 1936

Statistics replaced geodesy as the main use for linear algebra

- continental surveys done
- data analysis important for agriculture and government

Karl Pearson

- led early computational developments
- *Tracts for Computers*

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

Calculators and Card Equipment in the 1920’s

Computing room at the U. S. Department of Agriculture

Outline

History of Gaussian Elimination

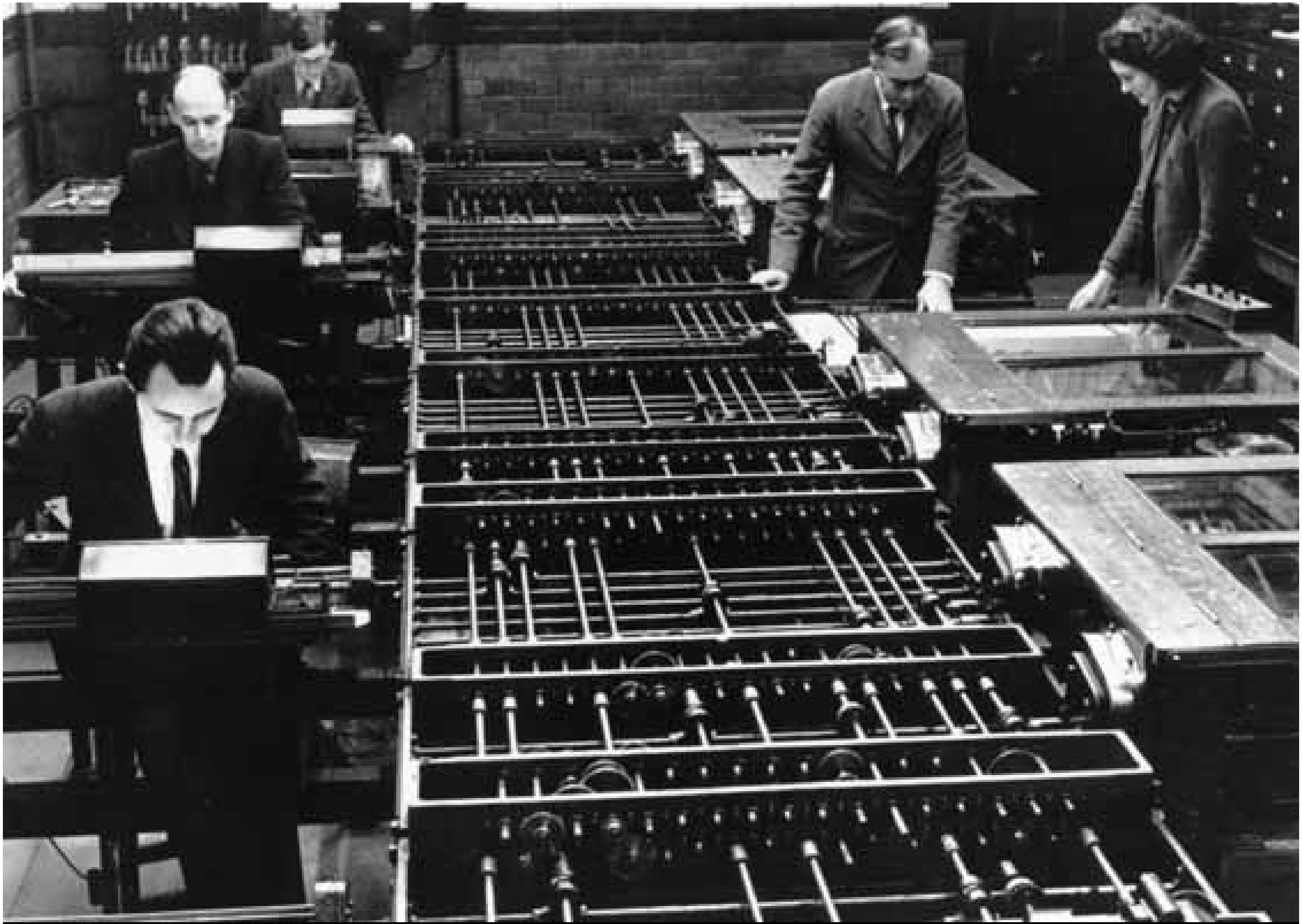
Invention of Electronic Computers

The Matrix Inversion Paper

Conclusion

Analog: MIT Differential Analyzers, 1931+

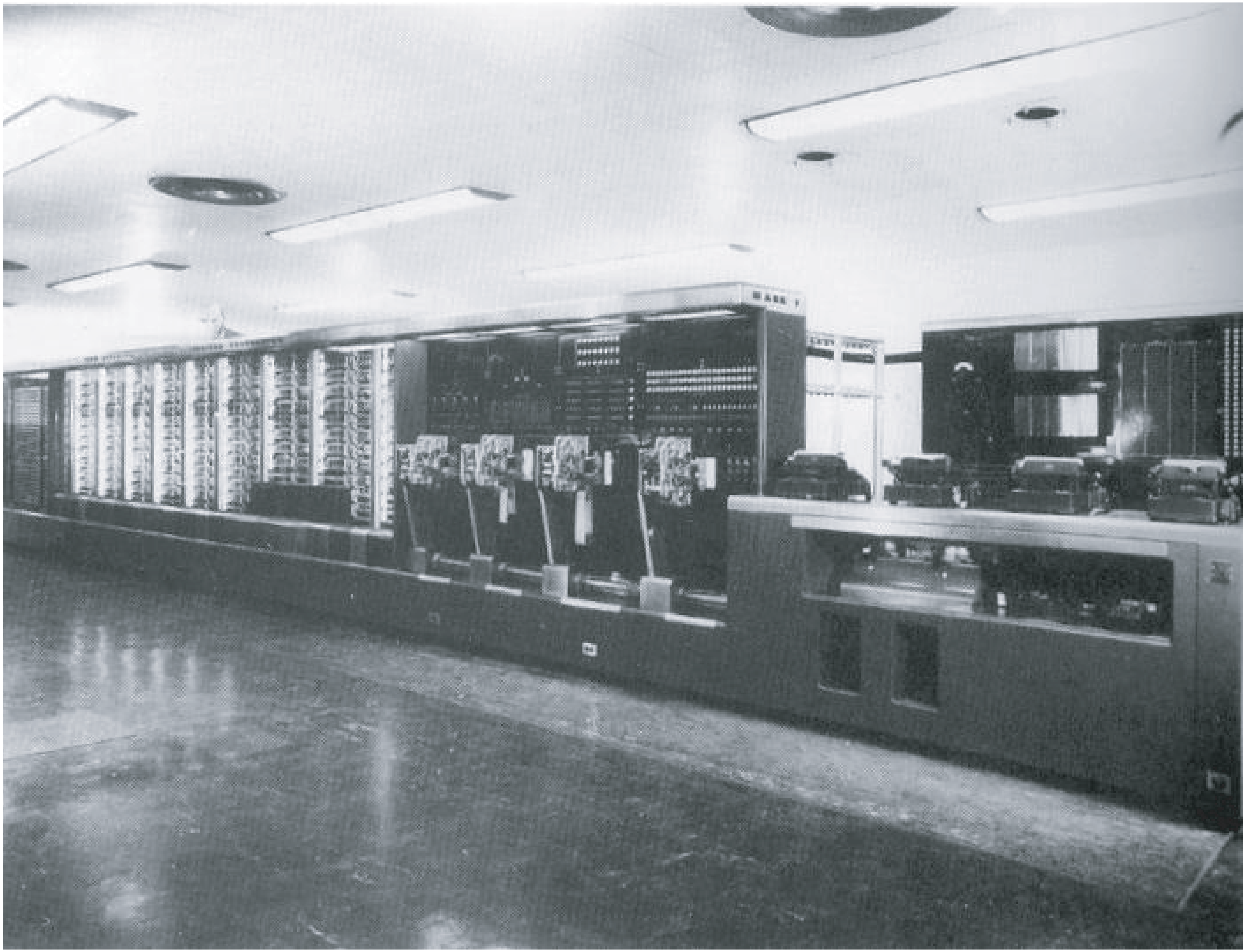
Analog: Manchester Differential Analyzer



Scan ©American Institute of Physics

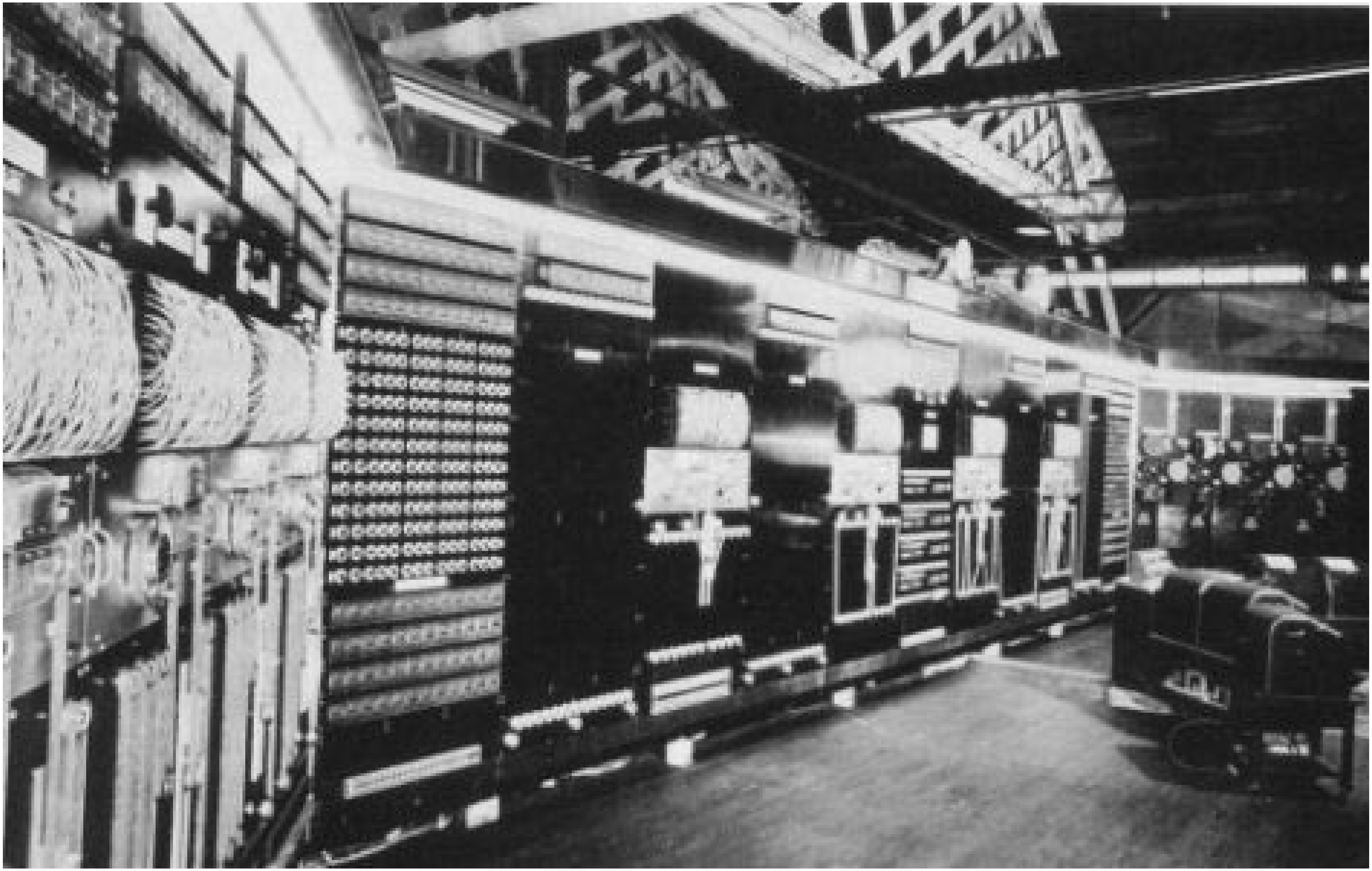
“Rockefeller” Differential Analyzer, 1942

Digital Mechanical: Harvard-IBM Mark I, 1944

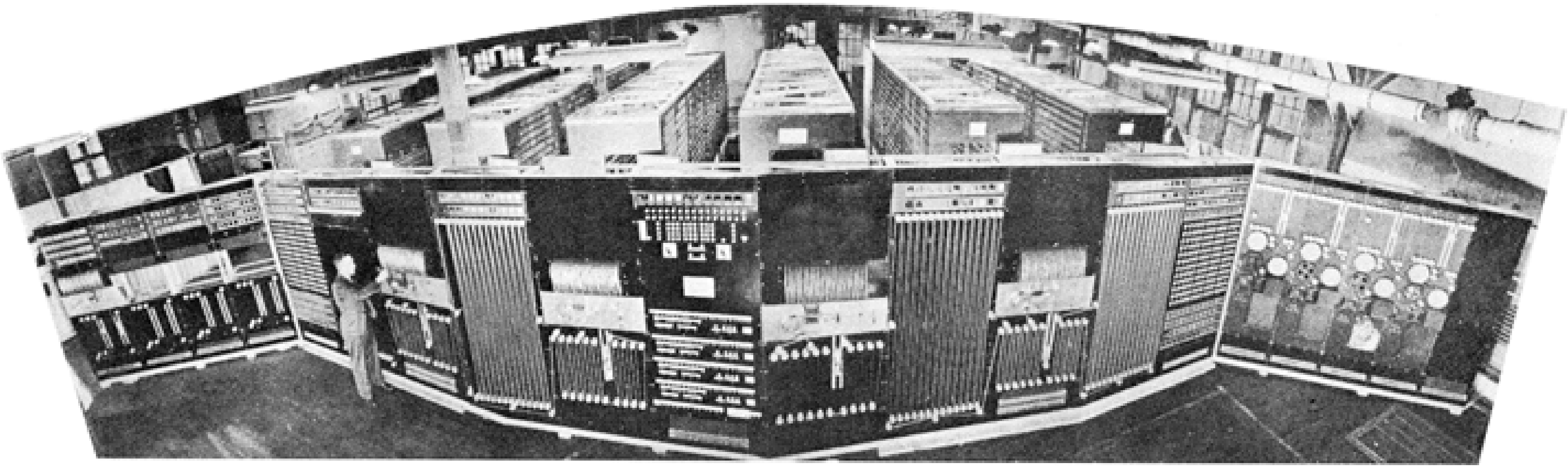


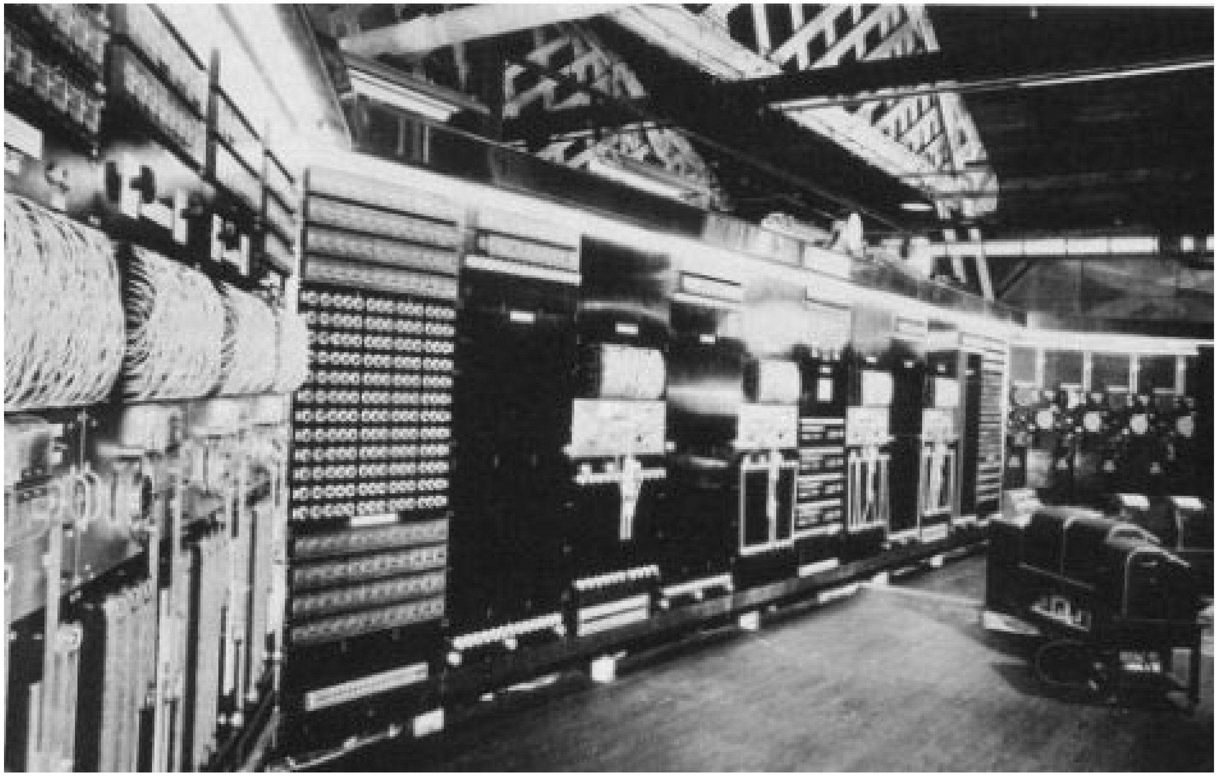
Electro-Mechanical: IBM SSEC, 1948

Electro-Mechanical: Harvard Mark II, 1947



Electro-Mechanical: Harvard Mark II, 1947





Analog

Electro-Mechanical Digital

MIT & NDRC

Vannevar Bush

1890 – 1974

Harvard

Howard Aiken

1900 – 1973

IBM

T. J. Watson Sr.

1907 – 1980

The "B" Team

University of Pennsylvania

John Presper Eckert

1919 – 1995

US Army Lieutenant

Herman Heine Goldstine

1913 – 2004

University of Pennsylvania

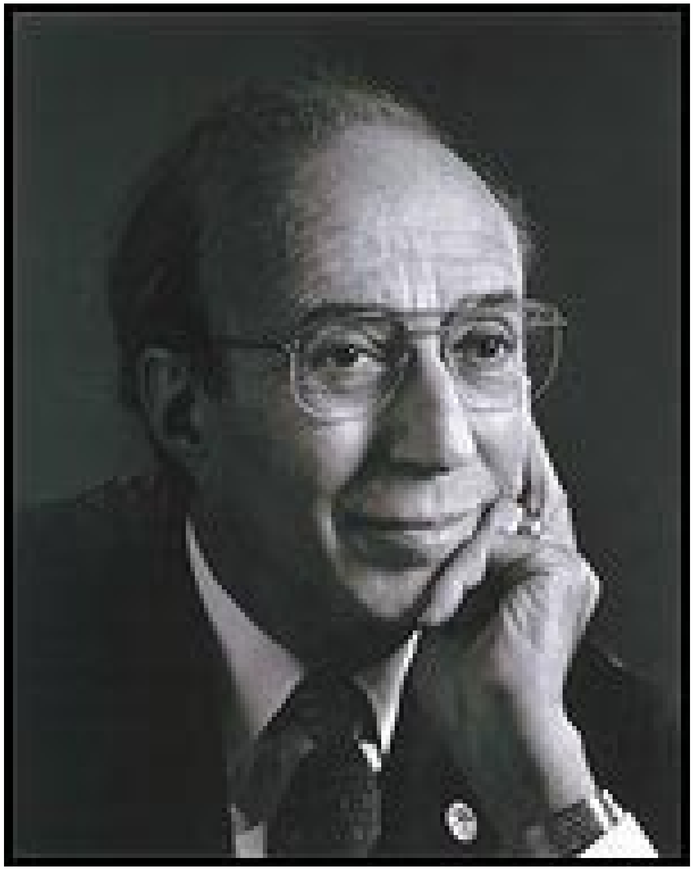
John William Mauchly

1907 – 1980

All-Electronic: Moore School ENIAC, 1946

"First Draft of a Report on the EDVAC"

Mimeographed distribution, June 1945



Herman Goldstine

1913 – 2004

John von Neumann

1903 – 1957

Moore School Lectures, Summer 1946

— Mina Rees, 1982

“Mauchly and Eckert had promised the machine in 18 months. . . . no matter when you asked them when would the machine be ready it was always in 18 months. They started calling that the von Neumann constant.”

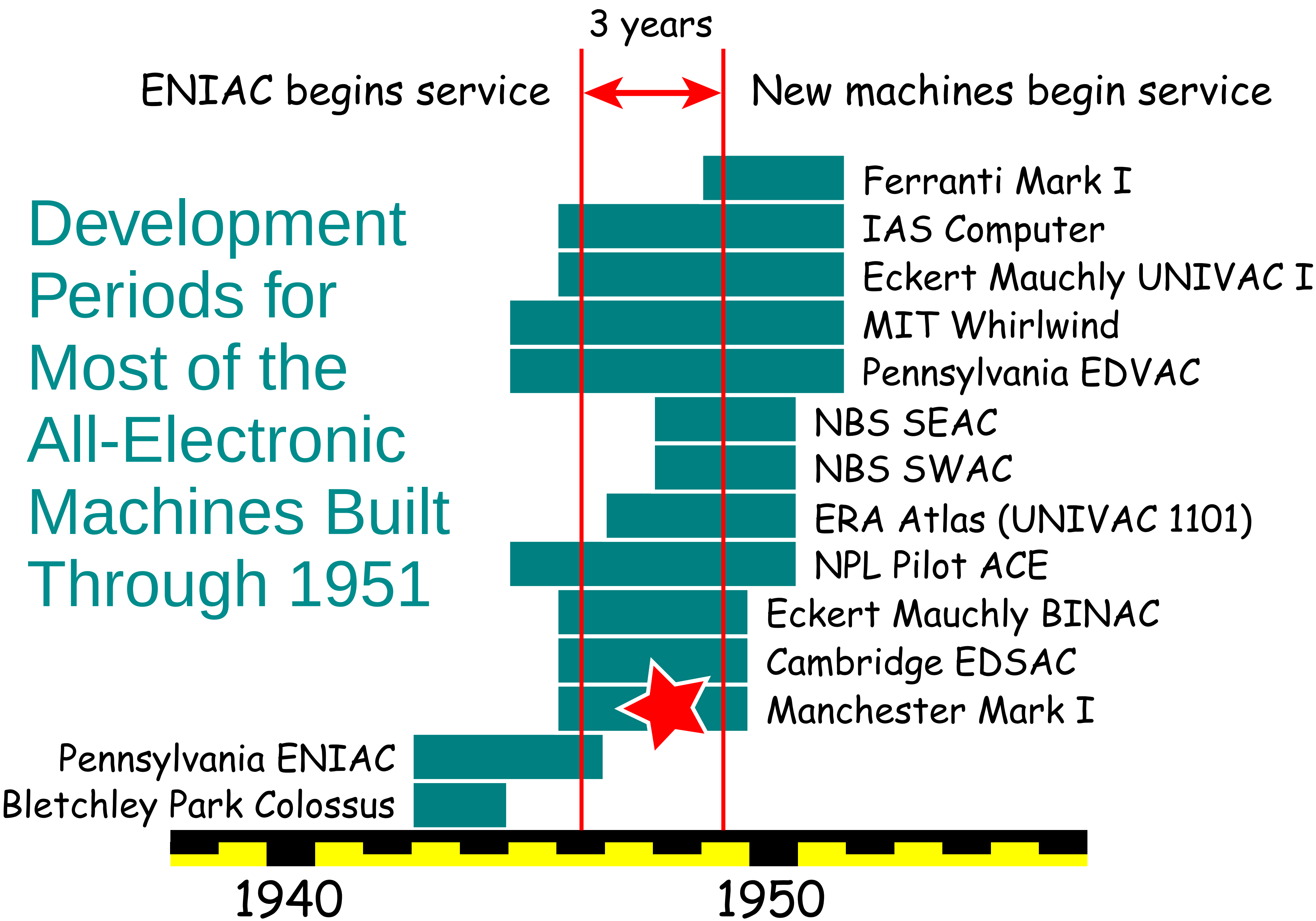
Smithsonian Oral History Interview, 1973

Ida Rhodes

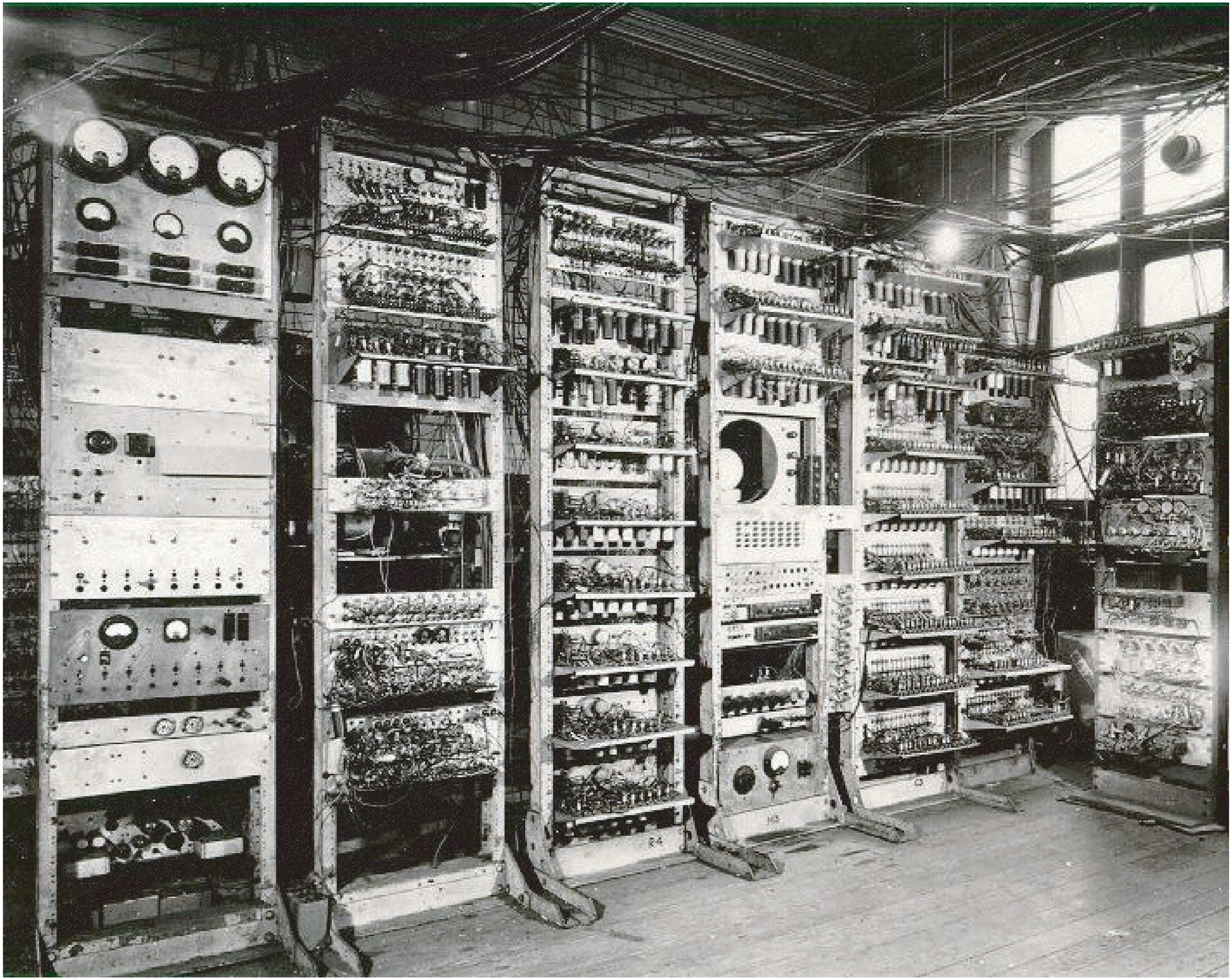
1900 – 1986

Years of the von Neumann Constant

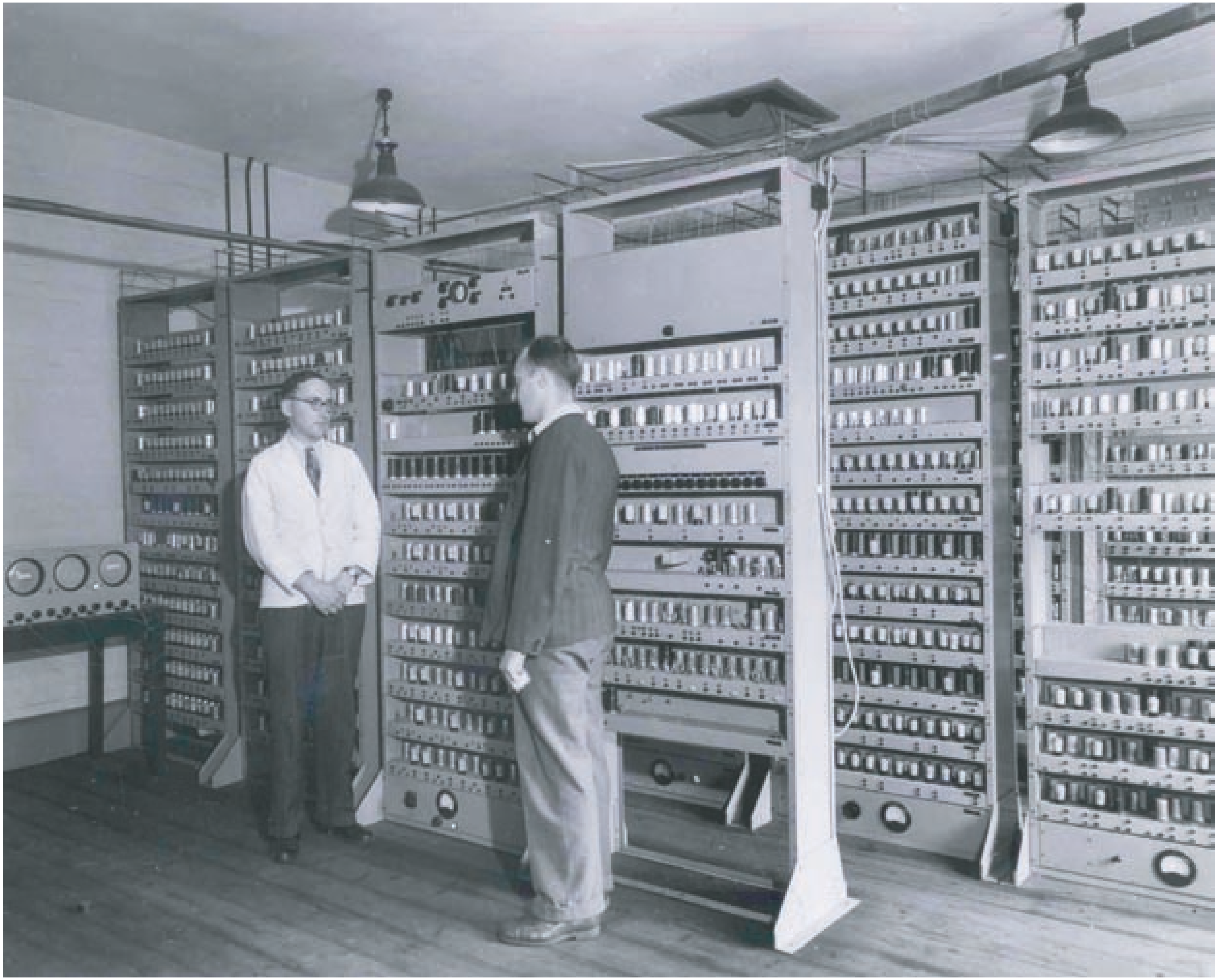
Development Periods for Most of the All-Electronic Machines Built Through 1951



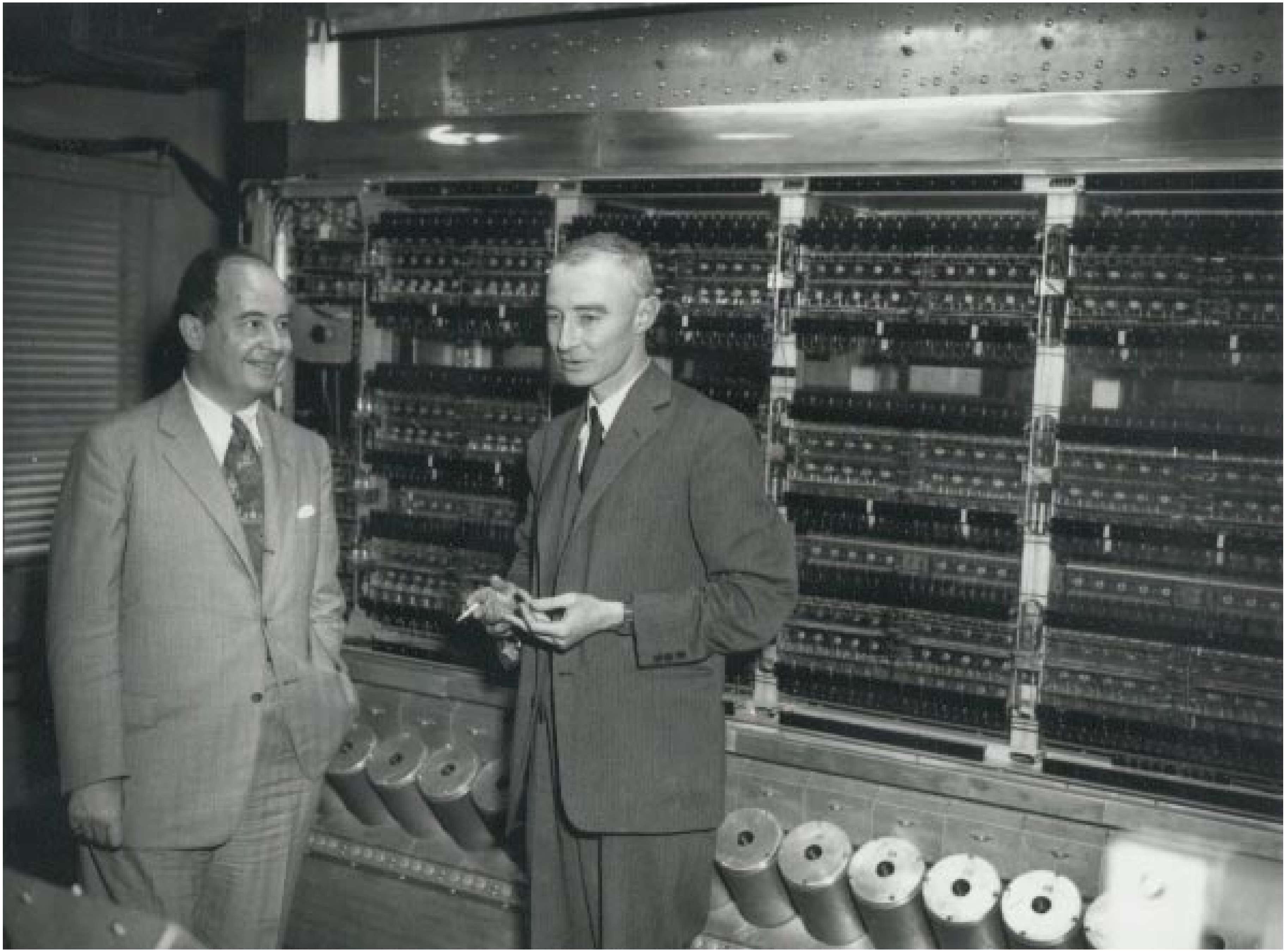
1st Computer: Manchester “Baby,” June 1948



2nd Computer: Cambridge EDSAC, May 1949

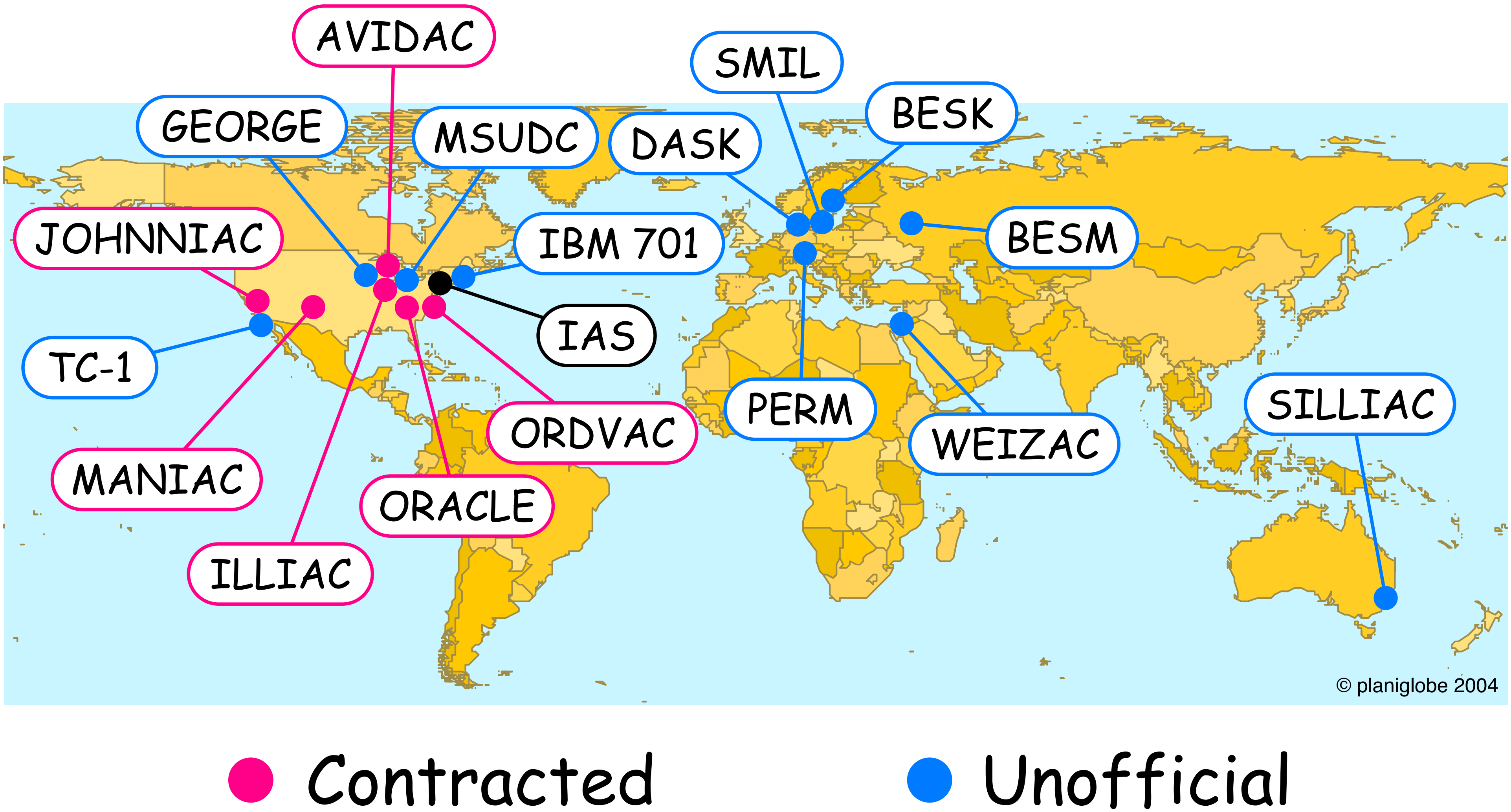


IAS Computer, 1951

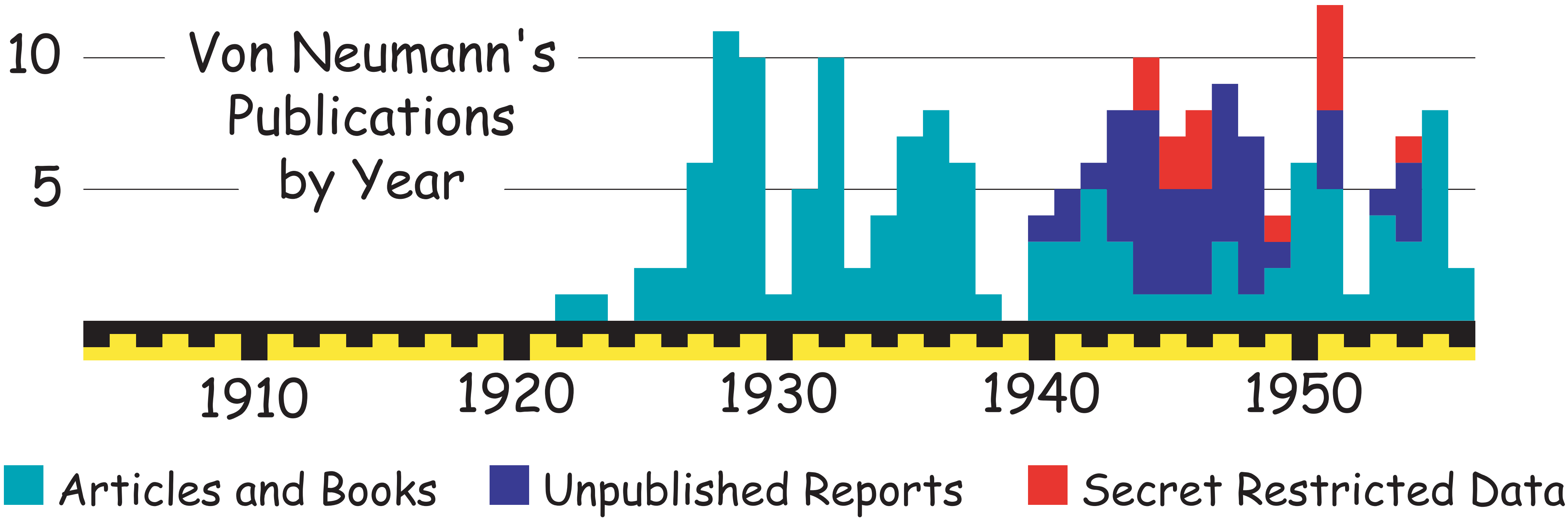


They Gave Away the Plans

Clones of the IAS Computer 1952–7



IAS Computer Reports and Plans



1945 *First Draft Report on the EDVAC*

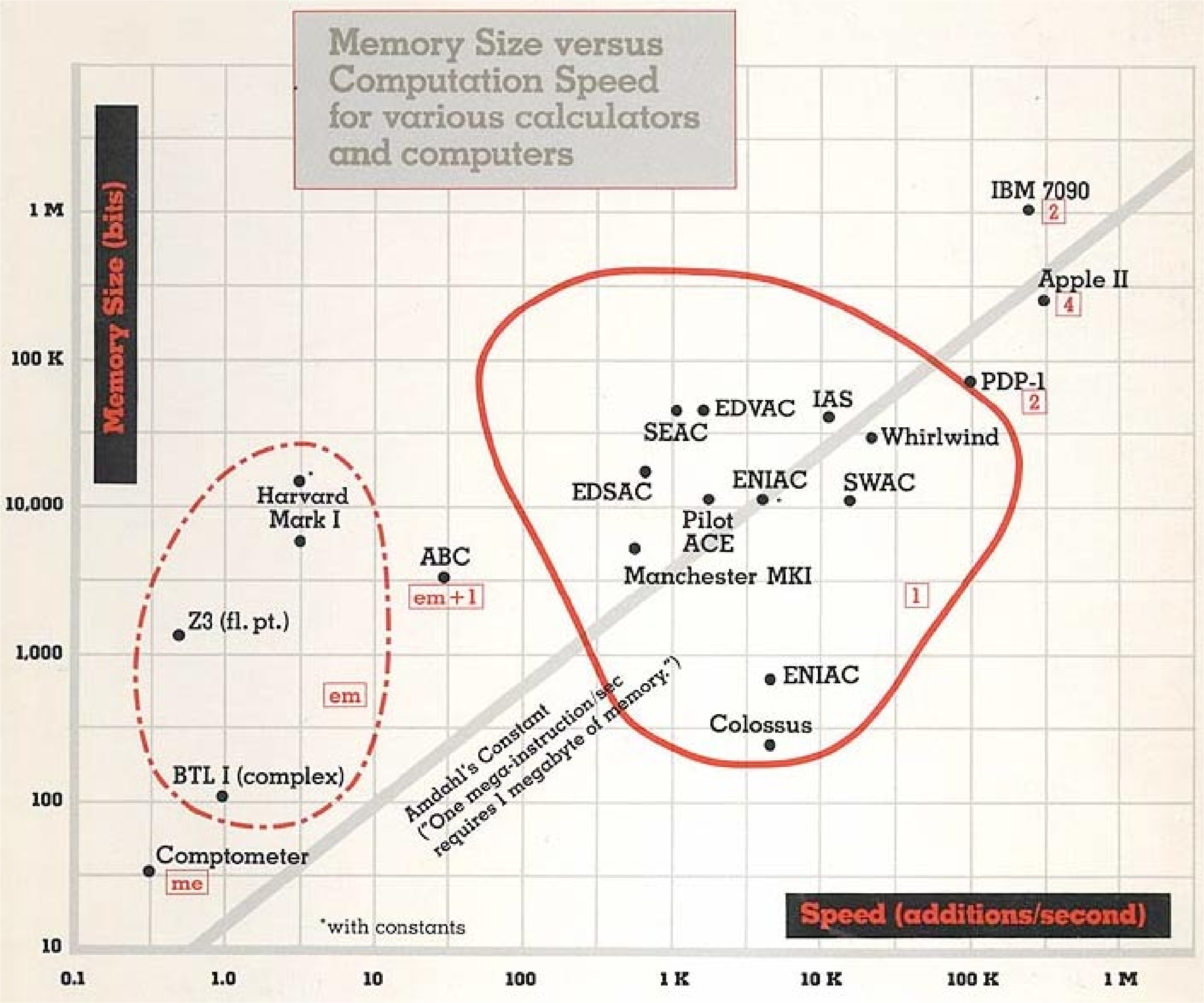
1946 *On the Principles of Large Scale Computing Machines*

1946 *Preliminary Discussion of the Logical Design of an Electronic Computing Instrument I.1*

1947 *Planning and Coding of Problems for an Electronic Computing Instrument II.1*

1948 *Planning and Coding of Problems for an Electronic Computing Instrument II.2*

Speed and Memory of Early Computers



"... approximation math-ematics will have to change very radically to use computers effectively — and to get into the position of being able to build still faster ones."

letter to Maxwell Newman, 1946

John von Neumann
1903 – 1957

Need for Mathematical Research

"...numerical analysis that would affect the design of computers and be responsive to the use of the powerful machines that, we were always confident, would one day come."

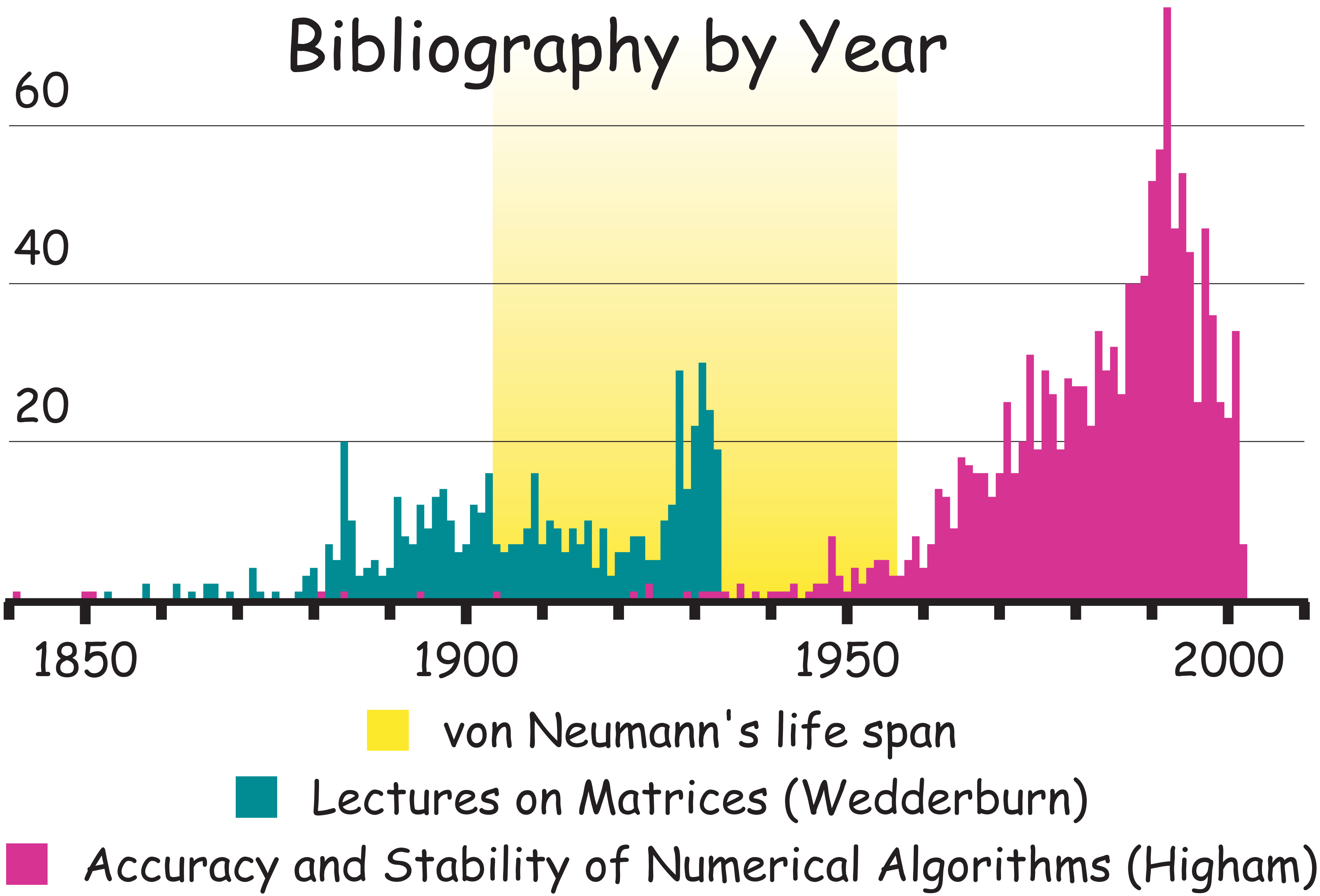
Annals of the History of Computing, 1982

Mina Rees
1902 – 1997
Director, Mathematics Branch, Office of Naval Research, 1946 – 1953

Outline

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

— Herman Goldstine, 1972



What Was the Proper Use of Computers?

"Should machines be used to generate tables of mathematical functions, or should they be used to solve problems directly in numerical form? There is a considerable difference of opinion . . ."

Moore School Lectures, 1946

George Stibitz
1904 – 1995

On the Accumulation of Errors

in Numerical Integration on the ENIAC

Rejected by *Mathematical Tables and Other Aids to Computation*

Hans Rademacher

1892 – 1969

Some New Methods in Matrix Calculation

The Annals of Mathematical Statistics

14(1):1–34, March 1943

Harold Hotelling

1895 – 1973

Numerical Inverting of Matrices of High Order

Bulletin of the American Mathematical Society

53(11):1021–1099, November 1947

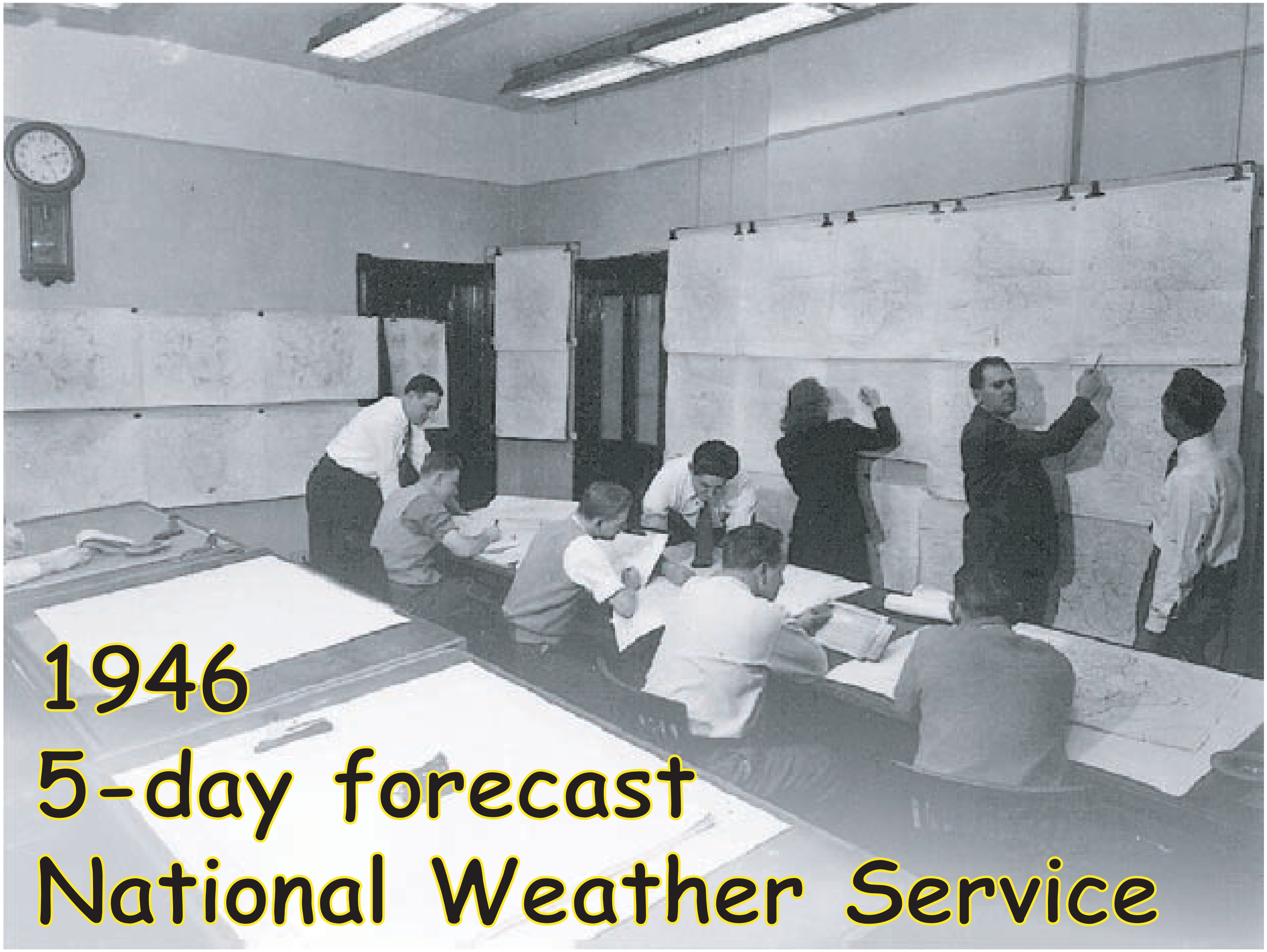
1. Introduced academic mathematicians to the idea of analyzing numerical algorithms.

2. Analysis was relevant because delays in building computers made experiments impossible.

3. The subject was topical because of prejudice that mechanized calculations might be error prone.

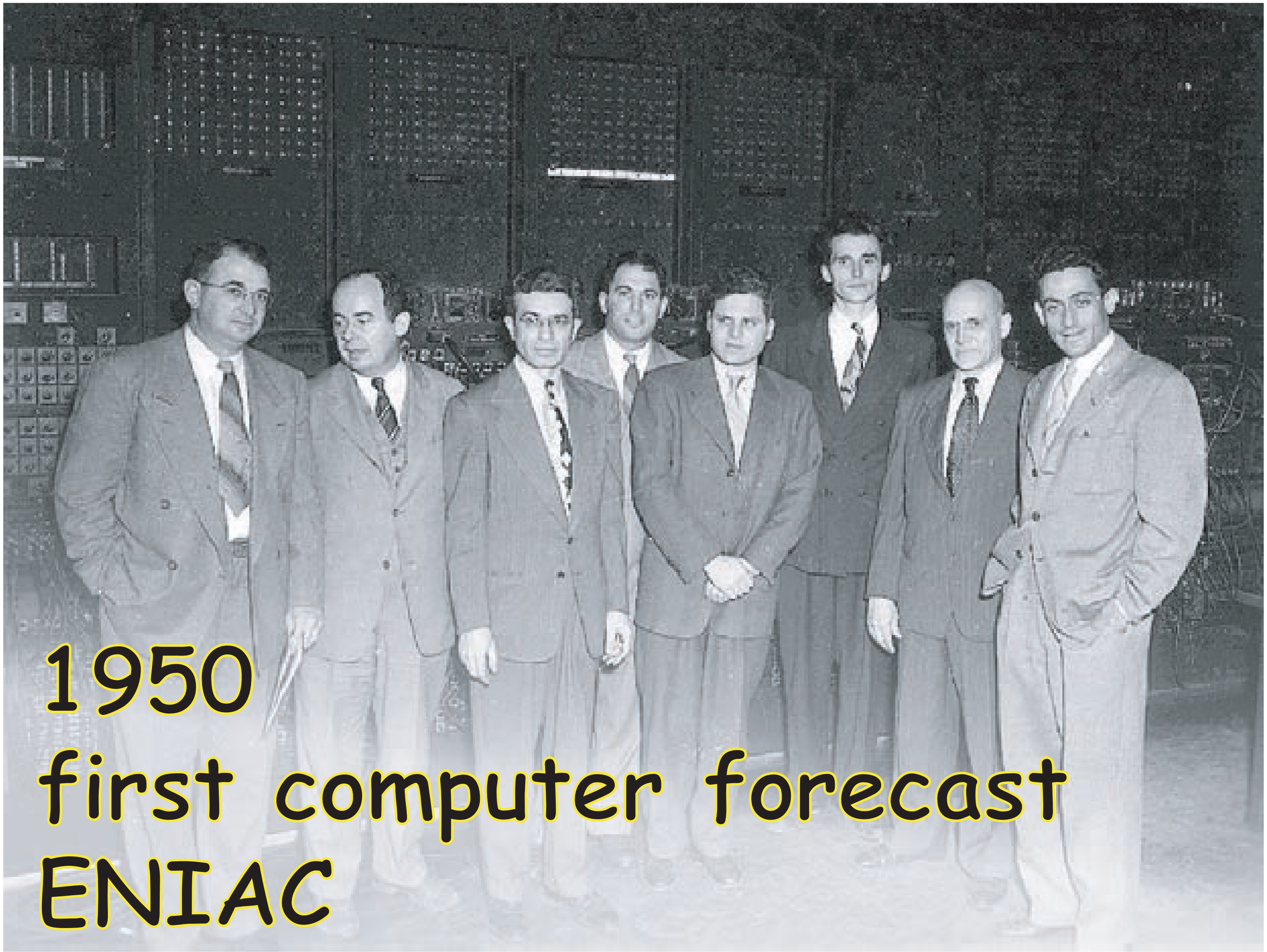
PDE's Were the Real Interest

Manhattan Project, Meteorology, Standard Oil . . .



PDE's Were the Real Interest

Manhattan Project, Meteorology, Standard Oil . . .



A^{-1} Was Not on the Critical Path

von Neumann knew elimination was irrelevant:

problem	size solved in 1960
$Ax = b$	100
PDEs	20,000

Venues of his matrix papers reflect his priorities:

1. “Some Matrix-Inequalities and Metrization of Matrix-Space,” **Tomsk. Univ. Rev.**, 1:286–300, 1937

2. “Numerical Inverting of Matrices of High Order,” **Bull. Amer. Math. Soc.**, 53:1021–1099, 1947

1st Modern Paper in Numerical Analysis

Numerical Inverting of Matrices of High Order

Bulletin of the American Mathematical Society

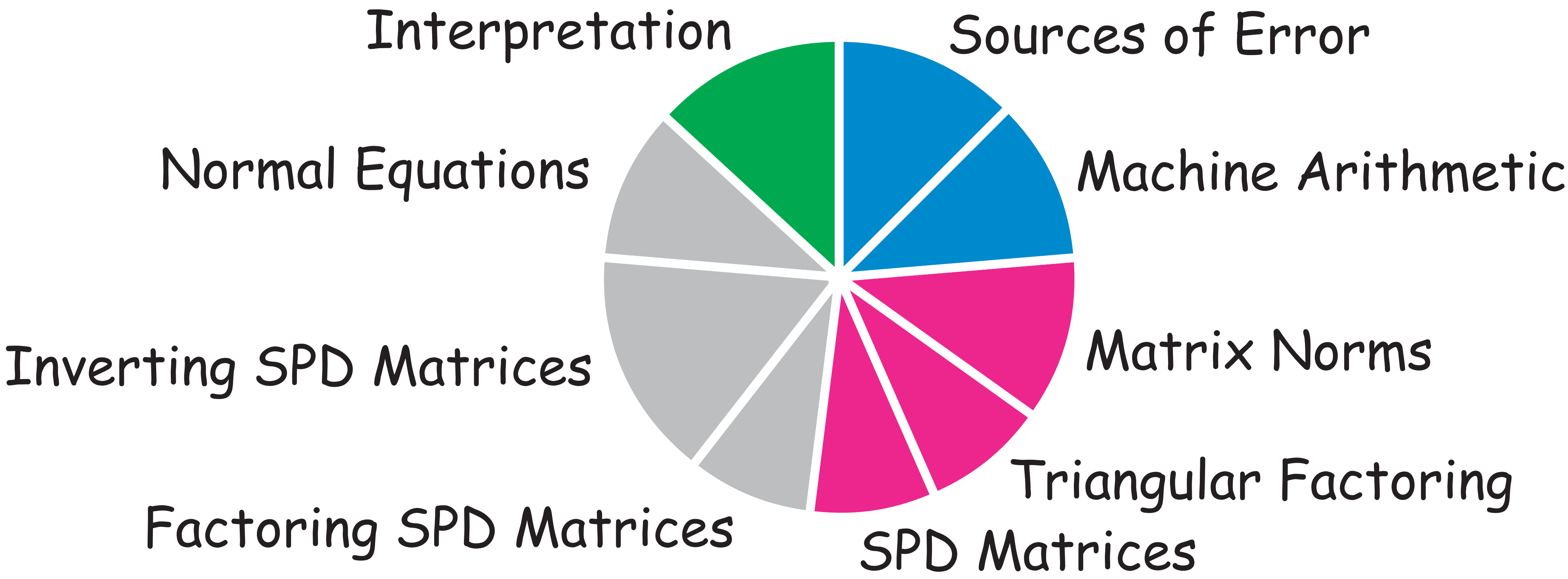
53(11):1021–1099, November 1947

1. Introduced academic mathematicians to the idea of analyzing numerical algorithms.
2. Analysis was relevant because delays in building computers made experiments impossible.
3. The subject was topical because of prejudice that mechanized calculations might be error prone.
4. Gaussian elimination was a convenient example otherwise unimportant to von Neumann.

"Get numerical analysis off to a good start"

— Herman Goldstine, 1972

Chapters of the Inversion Paper



“Get numerical analysis off to a good start”

— Herman Goldstine, 1972

1. Identified mathematical stability as a theme to unify the treatment of errors:

   ($\mathcal{A}$) Theory ($\mathcal{B}$) Measurement

   ($\mathcal{C}$) Discretization ($\mathcal{D}$) Rounding

2. Introduced the CFL condition.

3. Originated formal models of machine arithmetic:

$$\text{comp}(\bar{x} \bullet \bar{y}) = \text{round}(\bar{x} \bullet \bar{y})$$

4. Suggested mixing single and double precision.

5. Systematically described the use of matrix norms

"Get numerical analysis off to a good start"

— Herman Goldstine, 1972

6. Introduced the matrix lower bound.
7. Interpreted Gaussian elimination as triangular factoring, and discovered the LDU form.
8. First instance of the decomposition paradigm.
9. Discovered that SPD matrices have small "a priori" error bounds for factoring and inverting — the only significant matrices for which this is true.
10. Discovered the matrix condition number in error bounds for matrix inverting SPD matrices.

The Matrix Condition Number, ℓ

von Neumann to Goldstine from Bermuda, January 11, 1947

I have found a quite simple way to make the elimination method work. My results are these:

Consider a matrix A of order n, and with

$$\ell = \left(\frac{\text{largest proper value of } A^*A}{\text{smallest proper value of } A^*A}\right)^{\frac{1}{2}} = \frac{\underset{|f|=1}{\text{Max}} \; |Af|}{\underset{|f|=1}{\text{Min}} \; |Af|}.$$

Then if A is definite, the

“Get numerical analysis off to a good start”

— Herman Goldstine, 1972

11. Discovered that the condition number squared bounds the errors of the normal equations.

12. Inspired Alan Turing to invent the name condition number for measures of mathematical stability.

13. Introduced the concept of backward error.

14. Inspired Wallace Givens to invent backward error analysis.

15. Proposed “backward error vs measurement error” accuracy criterion rediscovered by Oettli & Prager.

— Herman Goldstine, 1972

16. Used stochastic linear algebra to estimate the size of matrices that could be inverted accurately.

$$B = \text{comp}\left[\bar{A}^t(\bar{A}\bar{A}^t)^{-1}\right], \quad \ell = \|A\|_2\,\|A^{-1}\|_2$$

$$\left\|\bar{A}\,B - I\right\|_2 \;\le\; \mathcal{O}\left(\ell^2\,n^2\beta^{-s}\right)$$

Bargmann asserted, and Edelman proved: if A has entries from a standard normal distribution, then the distribution for ℓ/n quickly approaches a limit for large n that peaks near 1.71.

$$\mathbf{Pr}\left[\ell/n \le 10\right] = 0.80$$

— Herman Goldstine, 1972

16. Used stochastic linear algebra to estimate the size of matrices that could be inverted accurately.

So $\mathbf{80\%}$ of the time the error bound is

$$\left\| \bar{A}\, B - I \right\|_2 \;\le\; \mathcal{O}(1)\, 10^2\, n^4 \beta^{-s}$$

For $\beta = 2$, $s = 40$, and $\mathcal{O}(1) = 14.24$,

$$\text{bound} \;\le\; 1 \quad \Longleftrightarrow \quad n \;\le\; 131$$

There was an $\mathbf{80\%}$ chance that the IAS computer could invert, with some accuracy, any randomly chosen matrix of order up to $\mathbf{131}$.

Outline

von Neumann and Goldstine's Contributions

1. Computers for computational science.
   ★ *They built the IAS computer.*

2. Independent research unencumbered by commercial interests or government secrecy.
   ★ *They gave away the plans.*

3. Mathematicians engaged in research about how computers can be used and improved.
   ★ *They wrote the inversion paper.*

Began DOE applied math program, 1956

Established the contract research program for mathematics and computer science at the AEC

John Pasta

1918 – 1981

John von Neumann

1903 – 1957

National Medal of Science

"For fundamental contributions to the development of the digital computer, computer programming and numerical analysis."

Ronald Reagan Library

John von Neumann's Analysis of Gaussian Elimination

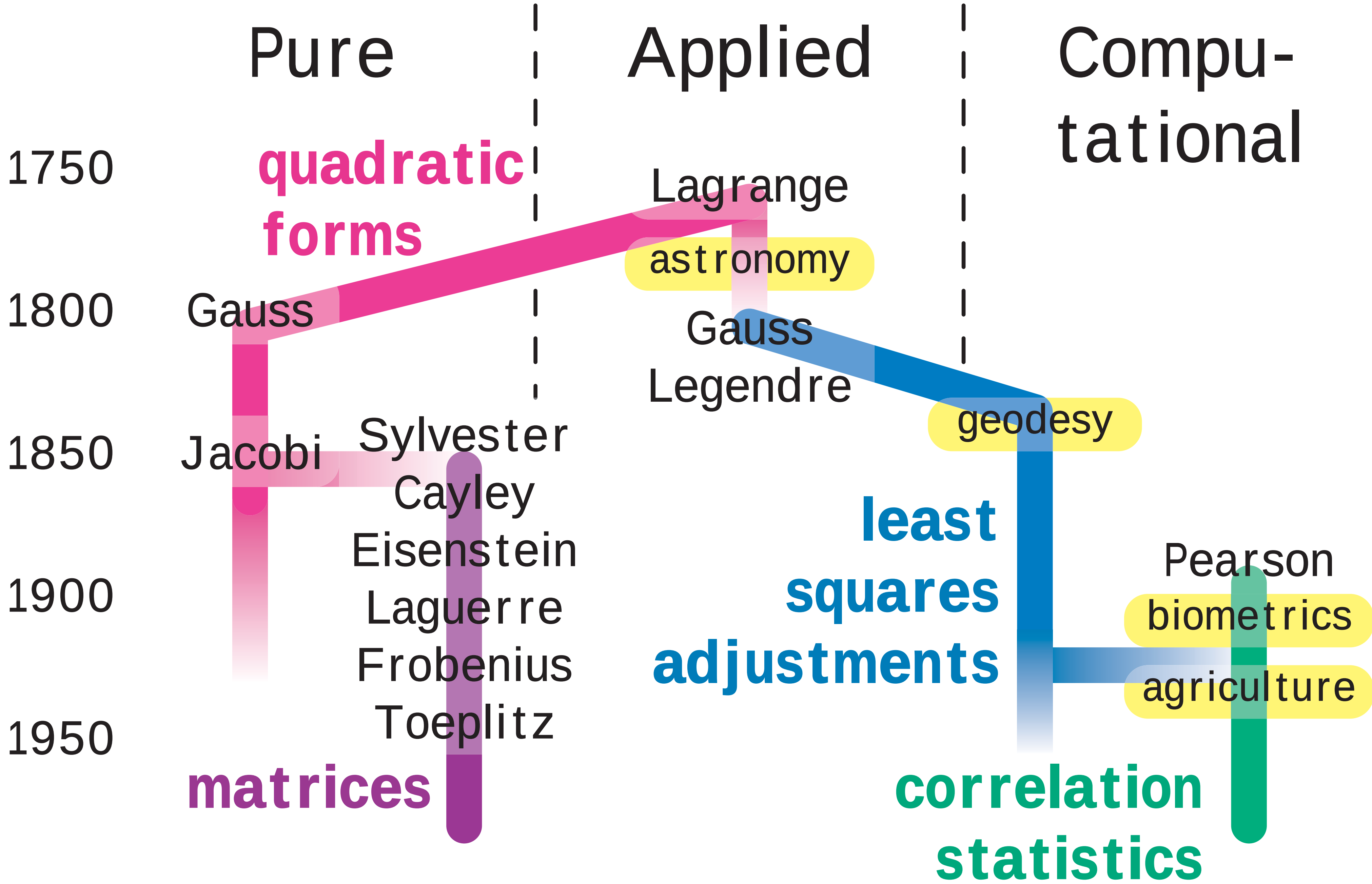
and the Invention of Modern Computing

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

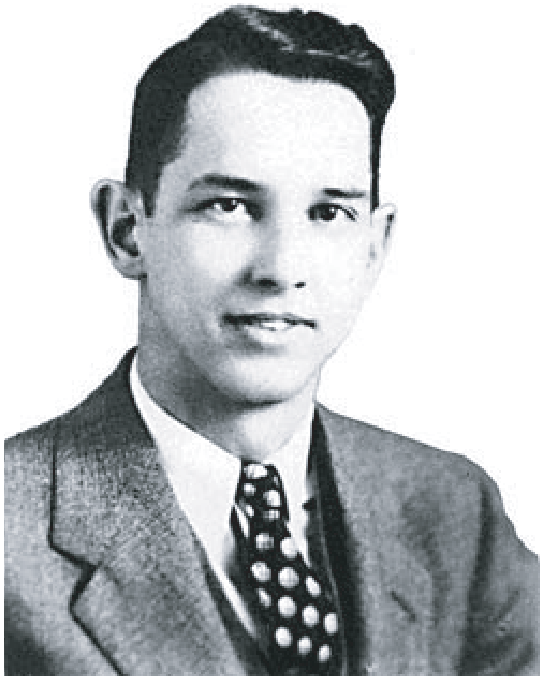
http://seesar.lbl.gov/ccse/people/grcar/Talks/talks.html

Extra Slides

Linear Algebra Timeline

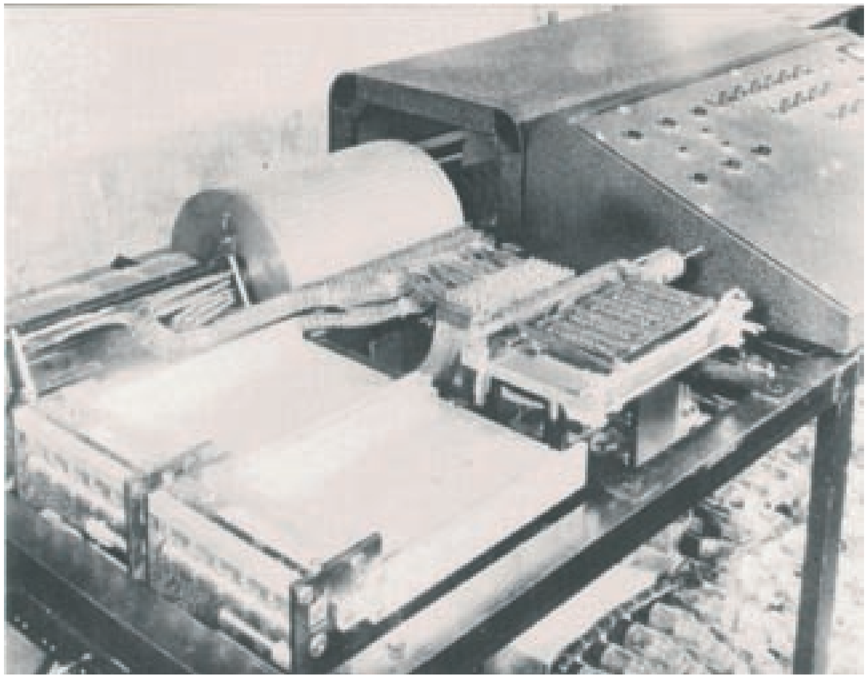


Iowa State University



John Vincent Atanasoff
1903 – 1995

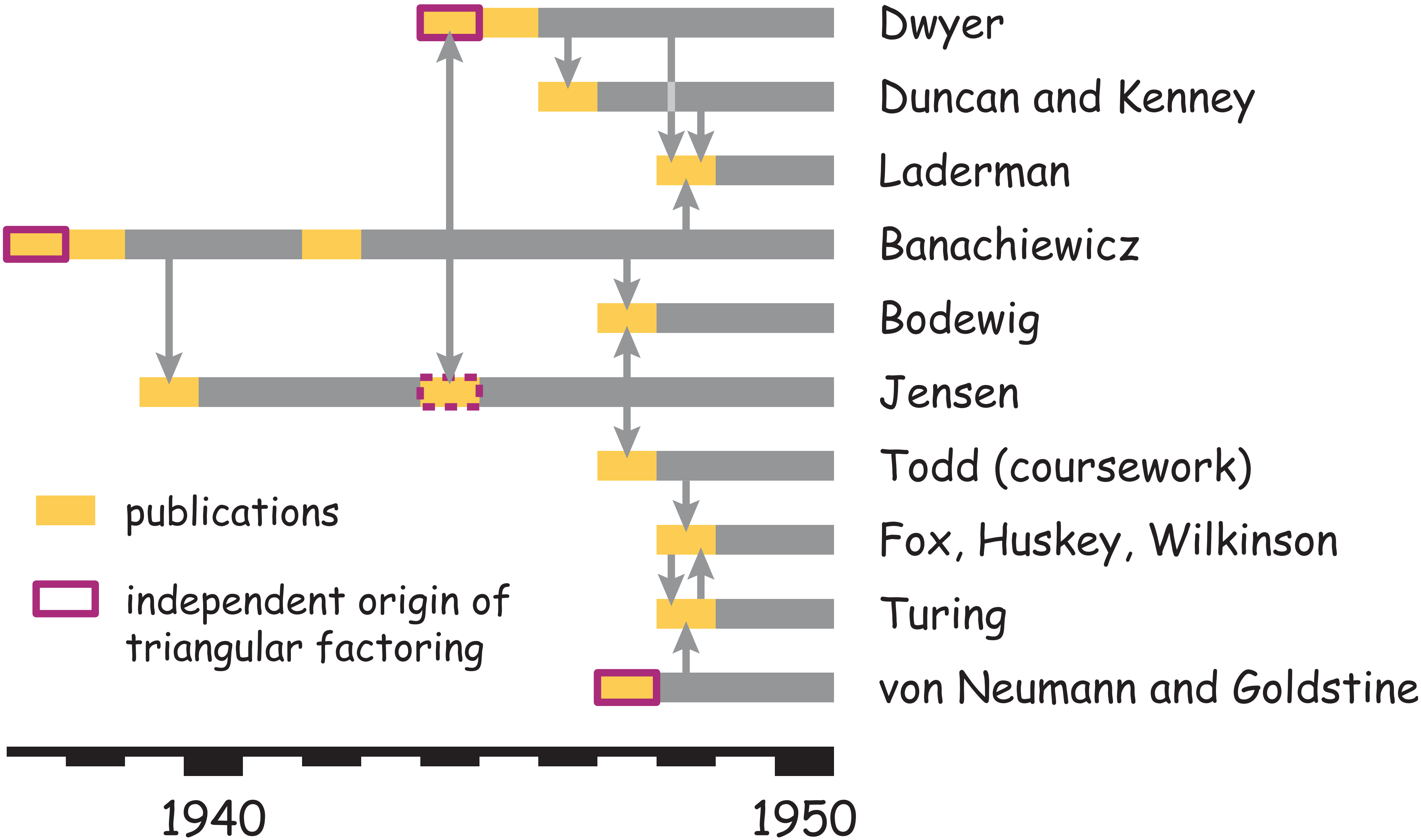
Clifford Edward Berry
1918 – 1963



Atanasoff-Berry Calculator
1942

Decomposition Paradigm

1st Papers Interpreting Gaussian Elimination as Triangular Factoring



Main Theorems of the Inversion Paper

	any A	SPD A
factoring	Gaussian elimination reordering $A = LDU$ unique L, D, U	Gaussian elimination symmetric reordering $\|\overline{A} - \overline{U}^t \overline{D}\,\overline{U}\|_2$ $\le \mathcal{O}(n^2 \beta^{-s})$
inverting	"normal equations" $A^{-1} = A^t (A A^t)^{-1}$ $\|\overline{A}\,\overline{W} 2^q - I\|_2$ $\le \mathcal{O}(\ell^2 n^2 \beta^{-s})$	product of inverses $A^{-1} = U^{-1} D^{-1} U^{-t}$ $\|\overline{A}\,\overline{W} 2^q - I\|_2$ $\le \mathcal{O}(\ell\, n^2 \beta^{-s})$

fixed point arithmetic of s digits in base β

$\ell =$ "figure of merit" or matrix condition number of $\overline{A}$

Matrix Condition Numbers

date	source	name	value
1942 1947	Lonseth	measure of sensitivity	$\sum_{i,j} \lvert \text{cofactor}_{i,j}(A)/\det(A) \rvert$ $= \sum_{i,j} \lvert (A^{-1})_{i,j} \rvert$
1947	von Neumann and Goldstine	figure of merit, ℓ	$\lVert A \rVert_2 / \lVert A \rVert_\ell$ $= \sigma_{\max}(A)/\sigma_{\min}(A)$ $= \lVert A \rVert_2 \, \lVert A^{-1} \rVert_2$
1948	Turing	N-condition nmbr	$\lVert A \rVert_F \, \lVert A^{-1} \rVert_F \,/\, n$
		M-condition nmbr	$n \max_{i,j} \lvert A_{i,j} \rvert \max_{i,j} \lvert A^{-1}_{i,j} \rvert$
1949	Todd	P-condition nmbr	$\max \lvert \lambda(A) \rvert / \min \lvert \lambda(A) \rvert$

John Todd began calling vN&G's value the condition number in 1958, but P- can be found until 1962.

"Those who cannot remember the past are condemned to repeat it."

George Sanatayana

REPORT TO THE PRESIDENT

Computational Science: Ensuring America's Competitiveness

President's Information Technology Advisory Committee

June 2005