# André Louis Cholesky: family, life and works

**Claude Brezinski**

Université des Sciences et Technologies de Lille
France

Let $A$ be a symmetric positive definite matrix.

It can be decomposed as $A = LL^T$ where $L$ is a lower triangular matrix with positive diagonal elements.

Then, the system $Ax = b$ writes $LL^T x = b$.

Setting $y = L^T x$, we have $Ly = b$.

Solving this lower triangular system gives the vector $y$.

Then $x$ is obtained as the solution of the upper triangular system $L^T x = y$.

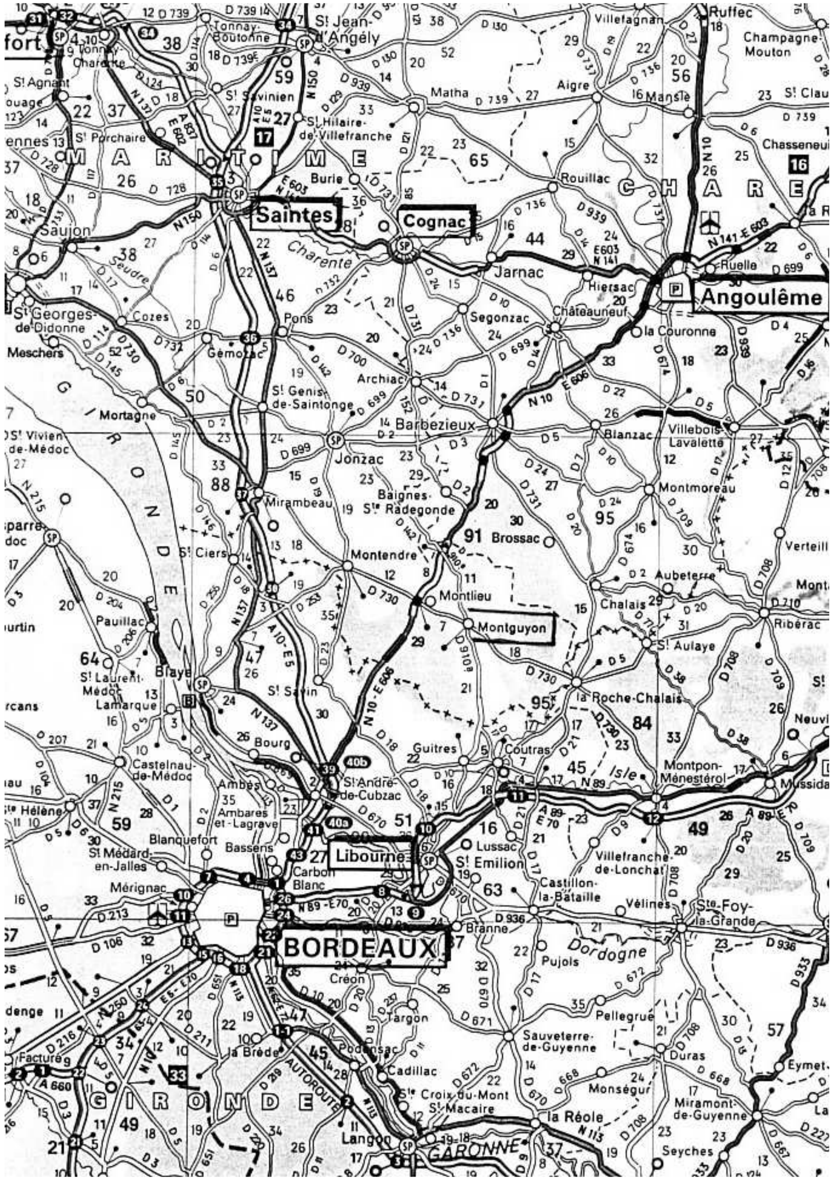
There are formulae for computing explicitly the elements of the matrix $L$.

This is **Cholesky's method**.

## The life of Cholesky

André Louis Cholesky was born on October 15, 1875 in Montguyon, around 35km nord–east of Bordeaux.

He was the son of André Cholesky, an hotel–keeper, and of Marie Garnier.


St Jean-d'Angély
Saintes
Cognac
Jarnac
Angoulême
Matha
Rouillac
Ruffec
Pons
Jonzac
Barbezieux
Archiac
Segonzac
Châteauneuf
Montmoreau
Brossac
Montendre
Montlieu
Montguyon
Chalais
Aubeterre
Ribérac
St Aulaye
la Roche-Chalais
Montpon-Ménestérol
Mussidan
Coutras
Guîtres
Mirambeau
St Ciers
Blaye
Bourg
St André-de-Cubzac
Libourne
St Emilion
Castillon-la-Bataille
Ste-Foy-la-Grande
BORDEAUX
Créon
Cadillac
Langon
la Réole
Sauveterre-de-Guyenne
Pauillac
Lesparre-Médoc
Castelnau-de-Médoc
Mérignac
Blanquefort
Royan
Saujon
Cozes
Mortagne
Meschers
Gémozac
St Genis-de-Saintonge
GIRONDE
CHARENTE
MARITIME
Dordogne
GARONNE
Isle
Charente


MONTGUYON (Charente-Inférieure)
Côté Ouest de la Place
HOTEL
DU CHEVAL BLANC


EDUCATION NATIONALE
COURS PROFESSIONNEL POLYVALENT RURAL
32
ASSOCIATION INTERMEDIAIRE V.I.E

He went to the high school in Saint-Jean-D′Angély.

He obtained his “baccalauréat” in Bordeaux in 1893.

In October 1895, he entered as a student at **École Polytechnique** for 2 years.

His professors were Camille Jordan and Georges Humbert for analysis, Émile Haag for geometry, Octave Callandreau for astronomy and geodesy, and Henri Becquerel for physics.

In October 1897, after finishing École Polytechnique, he became Sous-Lieutenant, and studied at the **École d'Application de l'Artillerie et du Génie** in Fontainebleau.

Among others, he had lectures on ballistic and topography.

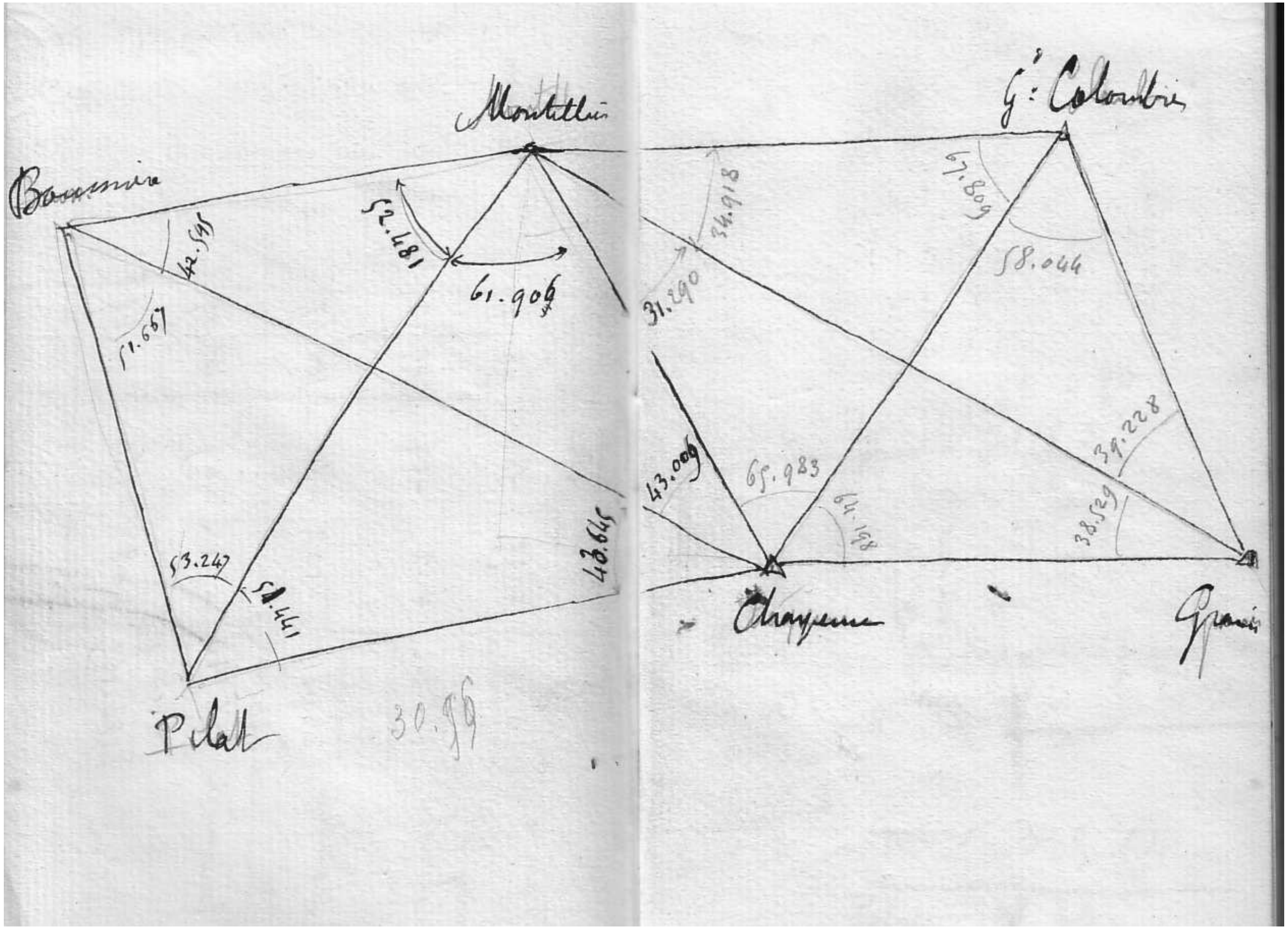
He ended 5th over 86.

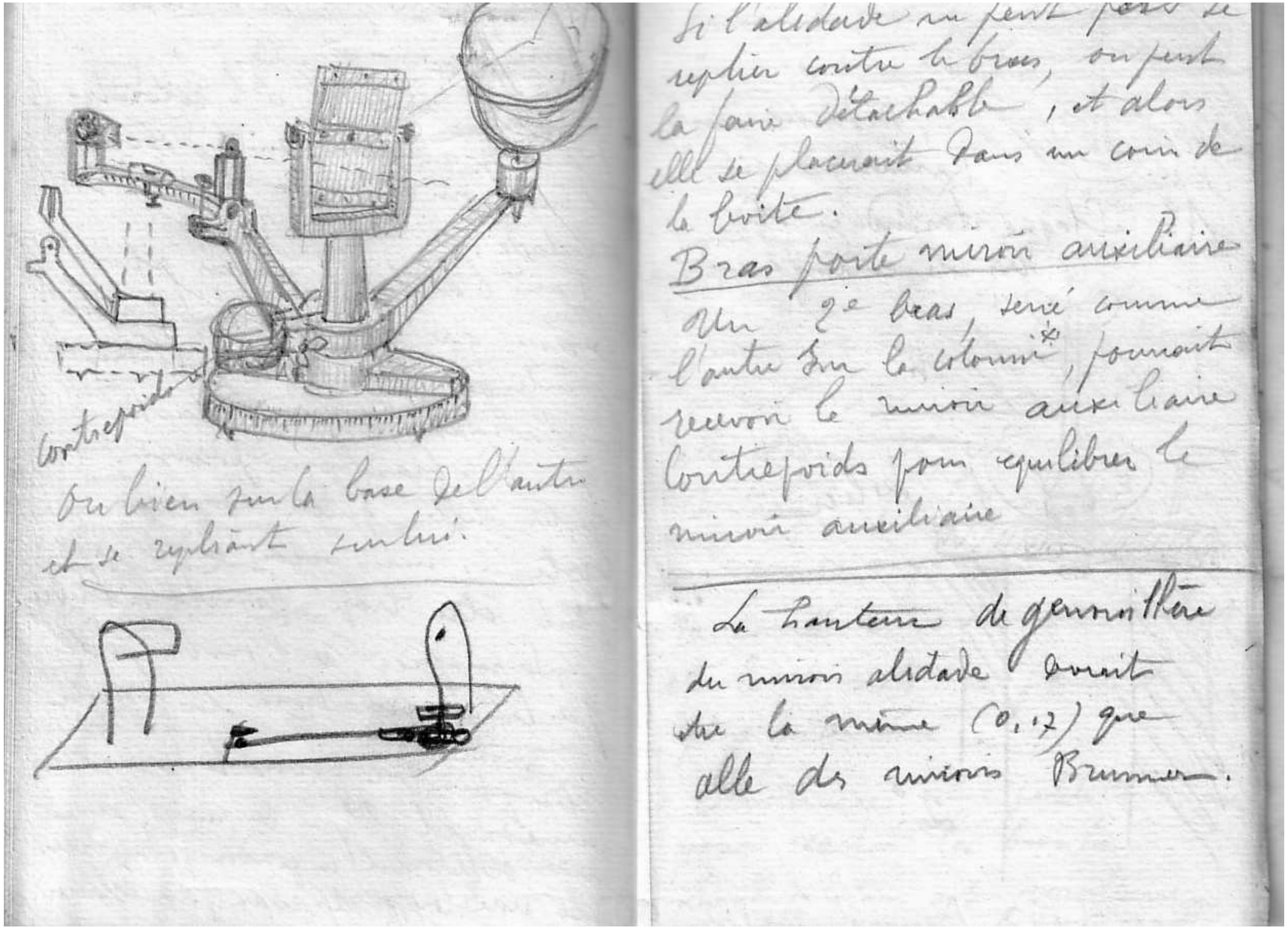
On October 1st, 1899, he became Lieutenant in the 22th Régiment d'Artillerie.

From January 17 to June 27, 1902 he was sent to Tunisia and again from November 21, 1902 to May 1st, 1903. From December 1903 to June 1906, he was in Algeria.

On June 24, 1905, he was appointed to the Geographical Service of the Headquarters of the Army. He was in the section having to mesure the length of the meridian of Lyon.

On May 10, 1907 he married his first cousin Anne Henriette Brunet, born June 27, 1882. They will have 4 children.


Montthlin
Baumine
G: Colombier
52.481
61.906
42.595
51.667
31.290
34.918
609.67
58.044
43.006
65.983
64.168
39.228
38.529
53.247
51.441
43.645
Pilat
Charpenne
Gravier
30.96



Contrepoids

Ou bien sur la base de l'autre
et se repliant sur lui.

Si l'alidade ne peut pas se
replier contre le bras, on peut
la faire détachable, et alors
elle se placerait dans un coin de
la boîte.

Bras porte miroir auxiliaire

Un 2e bras, serré comme
l'autre sur la colonne*, pourrait
recevoir le miroir auxiliaire
Contrepoids pour équilibrer le
miroir auxiliaire

La hauteur de genouillère
du miroir alidade devrait
être la même (0,17) que
celle des miroirs Brunner.

Cholesky went to Crete, then occupied by the international troops, from November 7, 1907 to June 25, 1908. He had to cartography the British and French parts of the island in the Lasiti plateau.

On March 25, 1909, he was promoted to Captain.

On August 28, 1909, he had to serve as the head of a battery for two years.

It is during this period that he found his method for the solution of systems of linear equations.

In October 1911, he was again appointed to the Geographical Service of the army as head of surveying in Tunisia and Algeria. He stayed in these countries until August 2, 1914, the day of the mobilization. He was head of the Topographical Service in Tunis.

After 3 months as the head of a battery, he was appointed at the Geographical Service and worked on canvas for shooting.

From September 1916 to February 1918, he was sent to Romania and became head of the Geographical Service that he completely reorganized.

On July 7, 1917 he was promoted Commandant.

In June 1918, he was appointed at the 202th Regiment of Artillerie of the Army of General Mangin. This army was fighting in Picardie between the river Aisne and Saint-Gobain.

On August 31, 1918 Cholesky was killed north of Bagneux, around 10km north of Soissons. He was first buried in the cemetery Chevillecourt near Autrèches, 15km east of Soissons.

On October 24, 1921 his grave was transferred to the cemetery of Cuts, 10km south-east of Noyon.

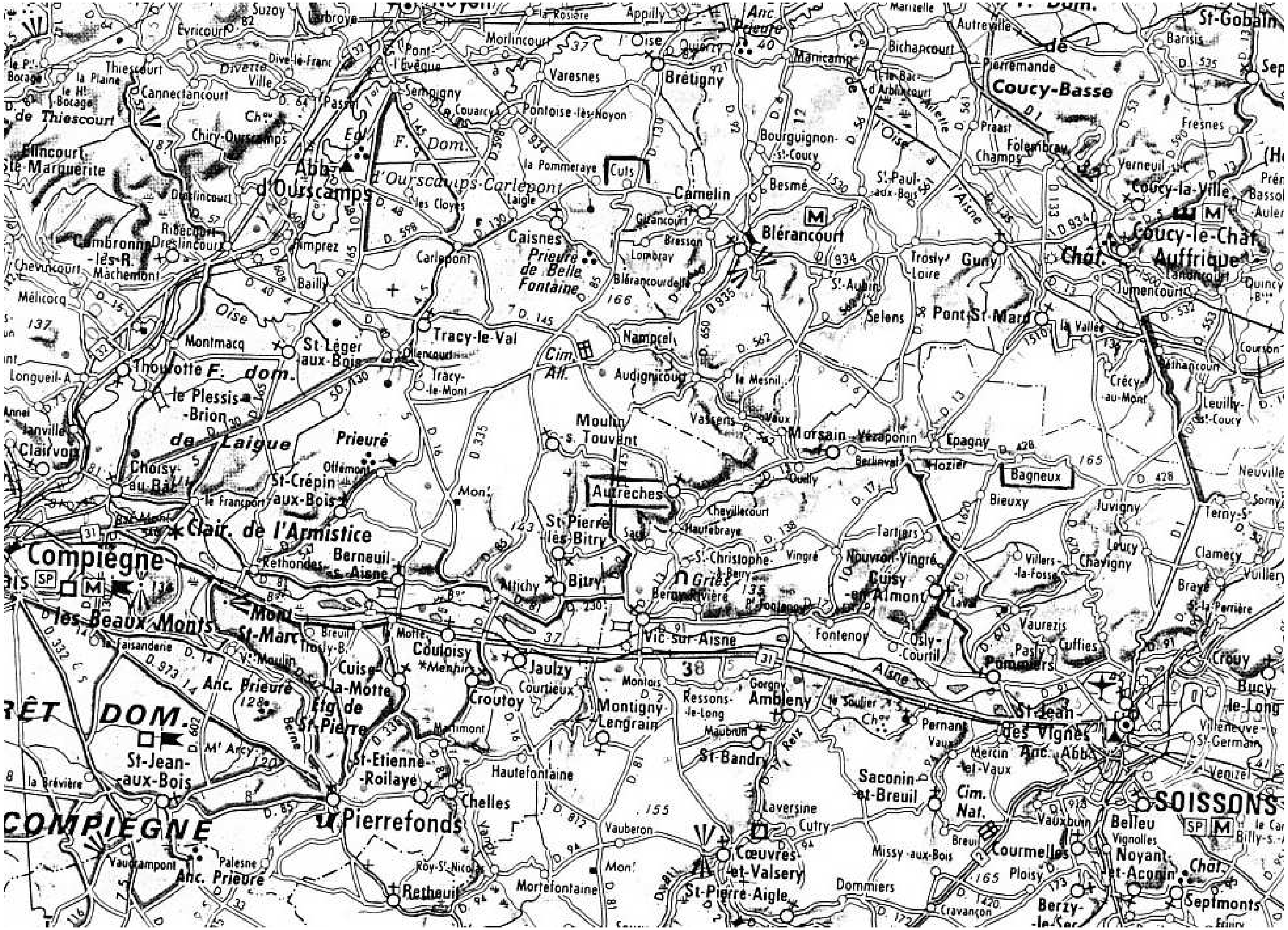
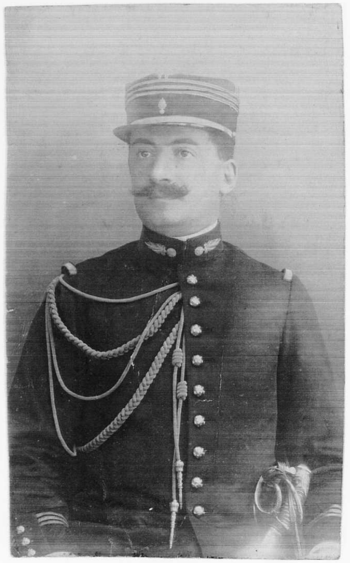



Figure 1: Bagneux, Autrèches and Cuts


CHOLESKY RENÉ
COMMANDANT 202e R.A.C.
MORT POUR LA FRANCE le 31-8-1918



## CHOLESKY'S SCIENTIFIC WORKS

Cholesky was a topograph in the Geographic Service of the Army.

His work mainly consisted in drawing maps of various scales.

From December 1909 and, at least, until January 1914 he was a Professor at the **École Spéciale des Travaux Publics, du Bâtiment et de l'Industrie** founded in 1891 by Léon Eyrolles. The studies were only by correspondence and Cholesky had to write lecture notes for the students.

In the archives, we found many manuscripts which, in fact, were incorporated in a book he published around these years. It contains 442 pages, 100 figures and 18 photos of instruments.

This book was quite successful since it had, at least, 7 editions and was still in the catalogue of the publisher 30 years after Cholesky's death.

The archives also contain another book with the title **Cours de Calcul Graphique**, several manuscripts describing the use of instruments for topography, and many military document on the organization of the work of a geographical officer, on shooting, etc.

ÉCOLE SPÉCIALE DES TRAVAUX PUBLICS

DU BATIMENT ET DE L'INDUSTRIE

M. LÉON EYROLLES, Ingénieur-Directeur.

# COURS DE TOPOGRAPHIE

2e PARTIE

## TOPOGRAPHIE GÉNÉRALE


COURS DE M. CHOLESKY

Ancien élève de l'École Polytechnique, Chef d'escadron d'artillerie,
Ancien Directeur des services topographiques de la Tunisie.

REVU PAR M. H. NOIREL

Répétiteur à l'École Polytechnique.


SEPTIEME EDITION

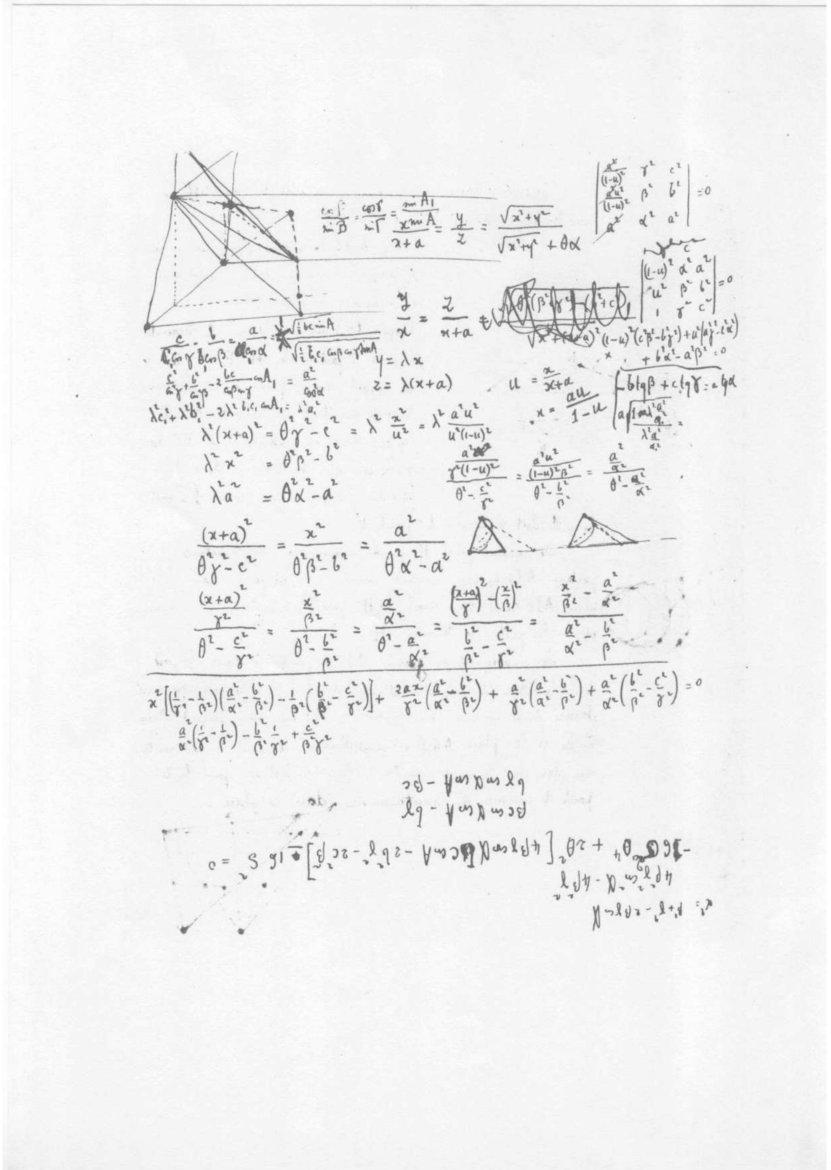

PARIS

ÉCOLE SPÉCIALE DES TRAVAUX PUBLICS

57, Boulevard Saint-Germain

PROPRIÉTÉ DU DIRECTEUR DE L'ÉCOLE

1937



## Historical context

A system of linear equations has infinitely many solutions when the number of unknowns is greater than the number of equations.

Among all possible solutions, one look for the solution minimizing the sum of the squares of the unknowns.

This is the case in the compensation of the triangles in topography which interested Cholesky.

sont déterminés successivement, les uns par les autres, à partir des sommets de la triangulation géodésique.

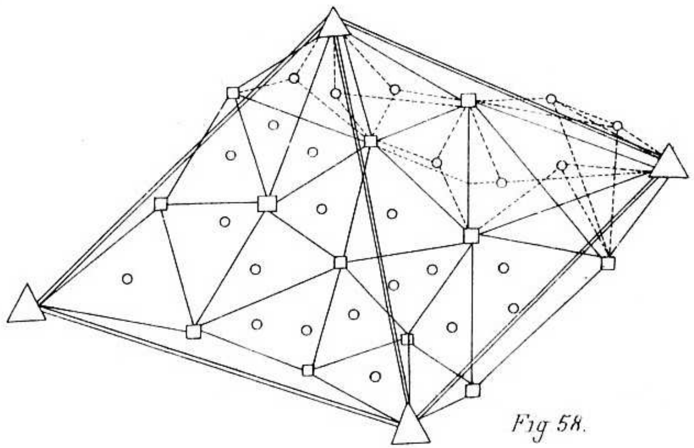


Fig 58.

△ Sommets de la Triangulation géodésique. ═══ Côtés de la triangulation géodesique

□ Points de $4^{me}$ ordre. ——— Visées qui déterminent le $4^e$ ordre

○ Points de $5^{me}$ ordre ------- Visées qui déterminent le $5^e$ ordre

On obtient ainsi une triangulation complémentaire comprenant des points de station déterminés directement à partir des signaux géodésiques par un enchaînement de triangles et d'autres points obtenus par intersections à partir des précédents. Comme les signaux géodésiques sont généralement répartis en Ier, 2ème et 3ème ordre, on désigne les points de la triangulation complémentai-

When the matrix $\mathbf{A}$ is **symmetric**, Gauss method makes no use of this property and needs too many arithmetical operations.

In 1907, **Otto Toeplitz** (Breslau, Allemagne, 1881 - Jerusalem, 1940) showed that an Hermitian matrix can be factorized into a product $\mathbf{L}\mathbf{L}^*$ with $\mathbf{L}$ lower triangular, but he gave no rule for obtaining the matrix $\mathbf{L}$.

This is what **Cholesky** did in 1910.

## AFTER CHOLESKY

Cholesky method was presented for the first time in 1924 in a note by **Commandant Benoît**, a French geodesist who knew Cholesky well.

But the method remained unknown outside the circle of French military topographers.

Cholesky method was rebirth by **John Todd** who taught it in his numerical analysis course at King's College in London in 1946 and thus made it known.

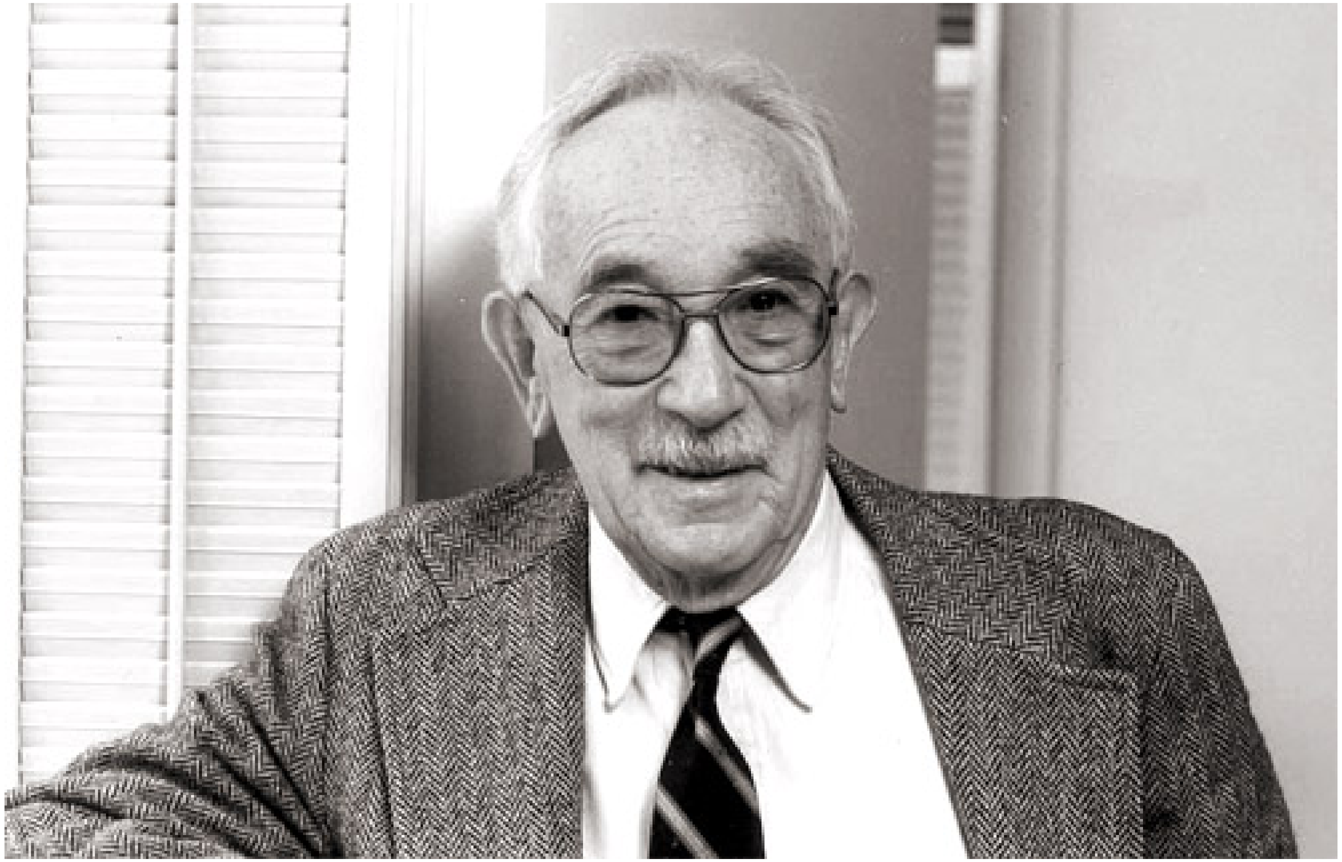
He tells

In 1946 one of us (J.T.) offered a course at Kings' College, London (KCL) on Numerical Mathematics. While we had some wartime experience in numerical mathematics, including characteristic values of matrices, we had had little to do with the solution of systems of linear equations. In order to see how this topic should be presented, we made a survey of Math. Rev. (at that time easy!) and found a review (MR7 (1944), 488), of a paper by Henry Jensen, written by E. Bodewig. Jensen stated "Cholesky's method seems to possess all advantages." So, it was decided to follow Cholesky and, since the method was clearly explained, we did not try to find the original paper.

John Todd continues

> **Leslie Fox (1918–1992), then in the newly formed Mathematics Division of the (British) National Physical Laboratory (NPL), audited the course and apparently found the Cholesky Method attractive, for he took it back to NPL, where he and his colleagues studied it deeply. From these papers, the Cholesky (or sometimes Choleski) Method made its way into the tool boxes of numerical linear algebraists via the textbooks of the 1950's.**

His colleagues were **James H. Wilkinson** and **Alan M. Turing**.



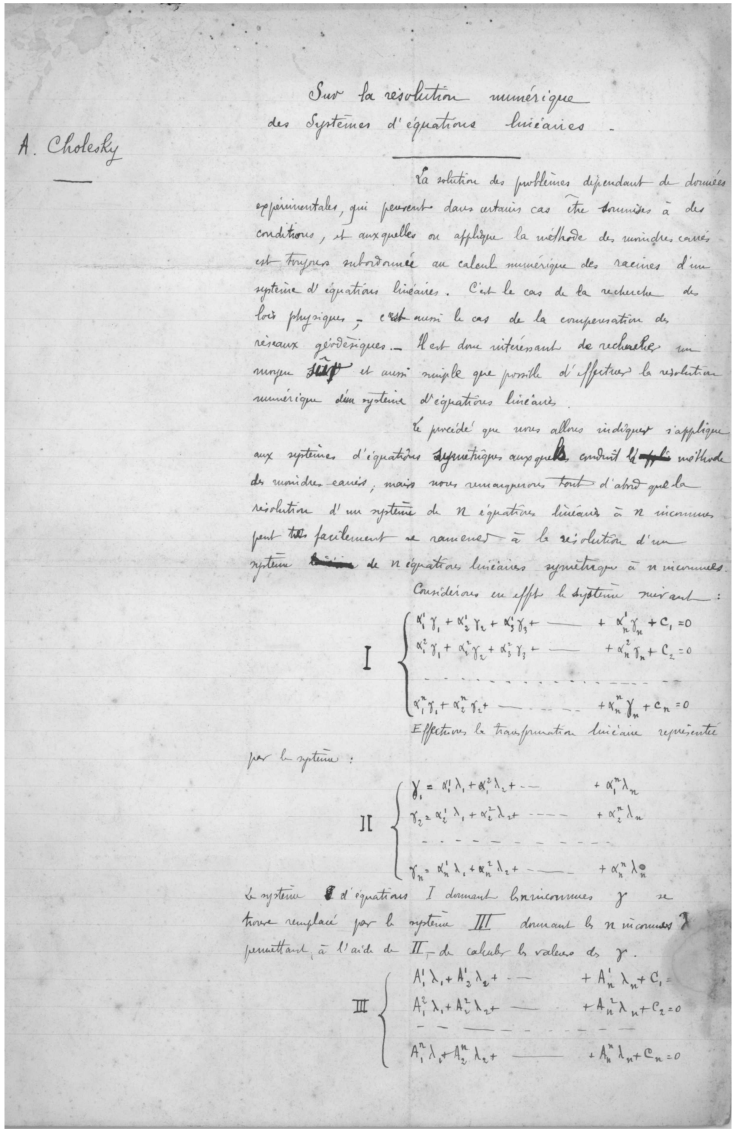
Some years ago, I was contacted by **Michel Gross**, one of Cholesky's grandsons.

The family wanted to give the documents they possessed to École polytechnique where Cholesky studied.

The grandson asked me if I wanted to help him to classify these papers.

I accepted, and we discovered an unpublished handwritten manuscript by Cholesky where he explained his method

# The manuscript



Sur la résolution numérique des Systèmes d'équations linéaires.

A. Cholesky

La solution des problèmes dépendant de données expérimentales, qui peuvent dans certains cas être soumises à des conditions, et auxquelles on applique la méthode des moindres carrés est toujours subordonnée au calcul numérique des racines d'un système d'équations linéaires. C'est le cas de la recherche des lois physiques ; c'est aussi le cas de la compensation des réseaux géodésiques. Il est donc intéressant de rechercher un moyen ~~sûr~~ et aussi simple que possible d'effectuer la résolution numérique d'un système d'équations linéaires.

Le procédé que nous allons indiquer s'applique aux systèmes d'équations ~~symétriques~~ auxquels conduit la ~~l'appli~~ méthode des moindres carrés ; mais nous remarquerons ~~tout~~ d'abord que la résolution d'un système de $n$ équations linéaires à $n$ inconnues peut ~~très~~ facilement se ramener à la résolution d'un système ~~linéaire~~ de $n$ équations linéaires symétrique à $n$ inconnues.

Considérons en effet le système suivant :

$$
\mathrm{I}\quad\begin{cases}
\alpha_1^1\gamma_1+\alpha_2^1\gamma_2+\alpha_3^1\gamma_3+ \text{——} \quad +\alpha_n^1\gamma_n+C_1=0\\
\alpha_1^2\gamma_1+\alpha_2^2\gamma_2+\alpha_3^2\gamma_3+ \text{——} \quad +\alpha_n^2\gamma_n+C_2=0\\
\text{- - - - - - - - - - - - - - -}\\
\alpha_1^n\gamma_1+\alpha_2^n\gamma_2+ \text{—— - - - -} \quad +\alpha_n^n\gamma_n+C_n=0
\end{cases}
$$

Effectuons la transformation linéaire représentée par le système :

$$
\mathrm{II}\quad\begin{cases}
\gamma_1=\alpha_1^1\lambda_1+\alpha_1^2\lambda_2+ \text{- -} \quad +\alpha_1^n\lambda_n\\
\gamma_2=\alpha_2^1\lambda_1+\alpha_2^2\lambda_2+ \text{- - - -} \quad +\alpha_2^n\lambda_n\\
\text{- - - - - - - - - - - -}\\
\gamma_n=\alpha_n^1\lambda_1+\alpha_n^2\lambda_2+ \text{- —— -} \quad +\alpha_n^n\lambda_n
\end{cases}
$$

Le système ~~I~~ d'équations I donnant les $n$ inconnues $\gamma$ se trouve remplacé par le système III donnant les $n$ inconnues $\lambda$ permettant, à l'aide de II, de calculer les valeurs des $\gamma$.

$$
\mathrm{III}\quad\begin{cases}
A_1^1\lambda_1+A_2^1\lambda_2+ \text{- —} \quad +A_n^1\lambda_n+C_1=\\
A_1^2\lambda_1+A_2^2\lambda_2+ \text{——} \quad +A_n^2\lambda_n+C_2=0\\
\text{- - —— - - - - - - -}\\
A_1^n\lambda_1+A_2^n\lambda_2+ \text{——} \quad +A_n^n\lambda_n+C_n=0
\end{cases}
$$

2)

On a d'une façon générale.

$$\text{IV)} \quad \begin{aligned} A_p^p &= \sum_{k=1}^{k=n} (\alpha_k^p)^2 \\ A_p^q &= \sum_{k=1}^{k=n} \alpha_k^p \alpha_k^q \end{aligned}$$

Le coefficient $A_p^q$ est obtenu en faisant le produit des coefficients de lignes $p$ et $q$ du système 1 qui se trouvent dans la même colonne et en faisant la somme des produits ainsi obtenus dans les $n$ colonnes, ce qu'on peut exprimer symboliquement en disant que $A_p^q$ est le produit de la ligne $p$ par la ligne $q$.

L'ordre des facteurs pouvant être renversé dans chaque produit, on voit immédiatement que

$$A_p^q = A_q^p$$

dans le déterminant du système III les termes symétriques par rapport à la diagonale sont égaux, autrement dit le système d'équations aux $\lambda$ est symétrique.

---

Proposons-nous donc de résoudre un système d'équations de la forme III.

Remarquons d'après ce qui précède que le système d'équations II si l'on y supposait les $\gamma$ connus serait un système d'équations aux $\lambda$ équivalent au système III. On aurait donc ~~immé~~ un moyen de résoudre le système III si l'on pouvait trouver un système I permettant de calculer facilement les $\gamma$.

C'est ce qui arrive si dans le système I la première équation ~~–~~ contient ~~que~~ seulement $\gamma_1$

la 2ème —————— $\gamma_1$ et $\gamma_2$

la 3ème $\gamma_1$, $\gamma_2$ et $\gamma_3$

ainsi de suite. On peut en effet calculer ainsi tous les $\gamma$ successivement à partir de $\gamma_1$.

Le problème est donc ramené à la recherche du système

$$\text{V)} \quad \begin{cases} \alpha_1^1\gamma_1 & + C_1 = 0 \\ \alpha_1^2\gamma_1 + \alpha_2^2\gamma_2 & + C_2 = 0 \\ \alpha_1^3\gamma_1 + \alpha_2^3\gamma_2 + \alpha_3^3\gamma_3 & + C_3 = 0 \\ - \; - \; - \; - \; - & \text{———} \\ \alpha_1^n\gamma_1 + \alpha_2^n\gamma_2 + \alpha_3^n\gamma_3 + \cdots & + \alpha_n^n\gamma_n + C_n = 0 \end{cases}$$

3)

Ce système étant en effet trouvé le problème devient ~~très~~ facile, puisque le système II est remplacé par le système VI qui permet de calculer les $\lambda$ de proche en proche à partir de $\lambda_n$.

$$
\text{VI)}\quad \begin{cases}
\alpha_1^1\lambda_1 + \alpha_1^2\lambda_2 + \cdots\cdots + \alpha_1^n\lambda_n - \gamma_1 = 0 \\
\qquad \alpha_2^2\lambda_2 + \alpha_2^3\lambda_3 \cdots\cdots + \alpha_2^n\lambda_n - \gamma_2 = 0 \\
\qquad\qquad \alpha_3^3\lambda_3 \cdots\cdots + \alpha_3^n\lambda_n - \gamma_3 = 0 \\
\qquad\qquad \cdots\cdots\cdots\cdots \\
\qquad\qquad\qquad \alpha_n^n\lambda_n - \gamma_n = 0
\end{cases}
$$

Nous calculerons facilement les coefficients $\alpha$ en partant des coefficients $A$ du système III, en appliquant les relations générales IV) au système V. On voit ainsi qu'on peut calculer ligne par ligne tous les coefficients du système VI

1ère ligne
$$
\begin{cases}
A_1^1 = (\alpha_1^1)^2 \\
A_2^1 = \alpha_1^1\alpha_1^2 \\
\cdots\cdots \\
A_p^1 = \alpha_1^1\alpha_1^p \\
A_n^1 = \alpha_1^1\alpha_1^n
\end{cases}
\qquad \text{d'où} \qquad
\begin{aligned}
\alpha_1^1 &= \sqrt{A_1^1} \\
\alpha_1^2 &= \frac{A_2^1}{\alpha_1^1} \\
\alpha_1^p &= \frac{A_p^1}{\alpha_1^1} \\
\alpha_1^n &= \frac{A_n^1}{\alpha_1^1}
\end{aligned}
$$

2ème ligne

les $\alpha_1$ sont déjà connus par le calcul de la 1ère ligne

$$
\begin{cases}
A_2^2 = (\alpha_1^2)^2 + (\alpha_2^2)^2 \\
A_3^2 = \alpha_1^2\alpha_1^3 + \alpha_2^2\alpha_2^3 \\
A_p^2 = \alpha_1^2\alpha_1^p + \alpha_2^2\alpha_2^p
\end{cases}
\qquad
\begin{aligned}
\alpha_2^2 &= \sqrt{A_2^2 - (\alpha_1^2)^2} \\
\alpha_2^3 &= \frac{A_3^2 - \alpha_1^2\alpha_1^3}{\alpha_2^2} \\
\alpha_2^p &= \frac{A_p^2 - \alpha_1^2\alpha_1^p}{\alpha_2^2}
\end{aligned}
$$

pème ligne

Tous les $\alpha$ dont l'indice inférieur est plus petit que $p$ sont connus par le calcul des lignes précédentes.

$$
\begin{cases}
A_p^p = (\alpha_1^p)^2 + (\alpha_2^p)^2 + (\alpha_3^p)^2 + \cdots\cdots + (\alpha_{p-1}^p)^2 + (\alpha_p^p)^2 \\
A_p^q = \alpha_1^p\alpha_1^q + \alpha_2^p\alpha_2^q + \alpha_3^p\alpha_3^q + \cdots\cdots + \alpha_{p-1}^p\alpha_{p-1}^q + \alpha_p^p\alpha_p^q \\
q > p
\end{cases}
$$

$$
\alpha_p^p = \sqrt{A_p^p - (\alpha_1^p)^2 - (\alpha_2^p)^2 - \cdots - (\alpha_{p-1}^p)^2}
$$

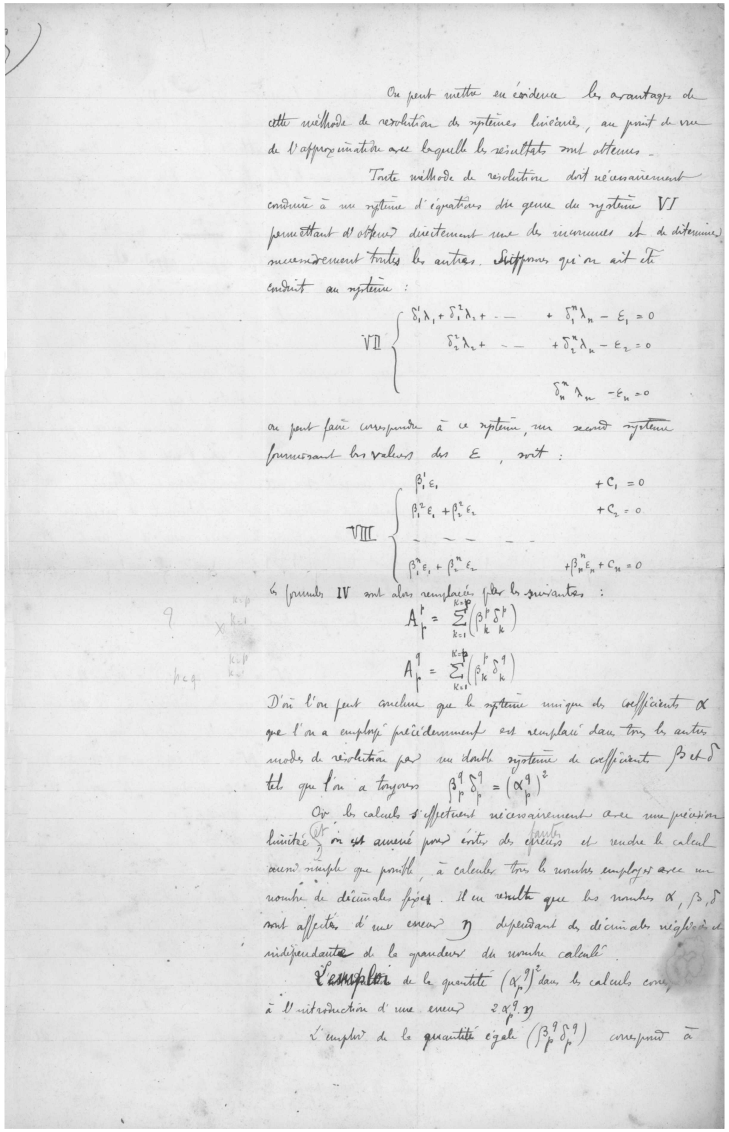
$$
\alpha_p^q = \frac{A_p^q - \alpha_1^p\alpha_1^q - \alpha_2^p\alpha_2^q - \cdots - \alpha_{p-1}^p\alpha_{p-1}^q}{\alpha_p^p}
$$

Quant au calcul des $\gamma$ il s'effectue facilement à l'aide des équations V. On obtient.

$$
\begin{aligned}
(-\gamma_1) &= \frac{C_1}{\alpha_1^1} \\
(-\gamma_2) &= \frac{C_2 - \alpha_1^2(-\gamma_1)}{\alpha_2^2} \\
&\cdots\cdots \\
(-\gamma_p) &= \frac{C_p - \alpha_1^p(-\gamma_1) - \alpha_2^p(-\gamma_2) \cdots}{\alpha_p^p}
\end{aligned}
$$

Ce qui montre que les coefficients $(-\gamma)$ qui figurent dans le tableau des équations VI) se calculent par rapport aux ~~coefficients~~ termes constants $C$ du tableau III exactement de la même façon que les coefficients $\alpha$ par rapport aux $A$.

Les calculs peuvent être disposés d'une façon
commode en un seul tableau. Les équations données étant
symétriques, il suffit d'écrire dans le tableau les coefficients
d'un seul côté de la diagonale, au-dessus par exemple.
Les équations transformées du système VI peuvent
alors être disposées au dessous de la diagonale symétriquement
placées par rapport aux équations données, chaque nouvelle
équation occupant une colonne au dessous de la diagonale.
Les écritures se bornent à la transcription des coefficients $\alpha$,
en effet le calcul d'un coefficient $\alpha$ de la forme $\frac{\Sigma mn}{K}$ est
effectué sur la machine à calculer à l'aide d'une
série de multiplications qui s'ajoutent automatiquement et
algébriquement sur la machine, la somme algébrique
étant immédiatement divisée par $K$. On utilise ainsi
l'opération la plus compliquée que puisse faire la machine
à calculer, par suite on obtient le rendement maximum
de cette machine, tout en se mettant autant que possible
à l'abri des fautes ~~erreurs~~ fréquentes dans les transcriptions de chiffres.
On trouve, tant dans la disposition des calculs,
que dans l'emploi de la machine à calculer, des simplifications
pour l'application des formules ou des indications précieuses.
Donnons un seul exemple : les machines à calculer du genre
« Dactyle » enregistrent le quotient [redacted] en chiffres blancs ou
rouges, suivant que la division [redacted] est faite en tournant
dans le sens de l'addition ou de la soustraction, il s'ensuit
[redacted] que tout coefficient $\alpha$ étant le résultat
d'une division sur la machine, son signe est
[redacted] indiqué par la couleur des chiffres qui le
composent ; les erreurs de signe sont aussi facilement évitées.
Il paraît inutile d'insister sur le reste de la
résolution, c'est à dire sur le calcul des $\lambda$ à l'aide
du système VI). On voit en effet immédiatement
comment on peut remonter successivement de $\lambda_n$
à $\lambda_{n-1}$, puis à $\lambda_{n-2}$ et ainsi de suite jusqu'à $\lambda_1$.



On peut mettre en évidence les avantages de cette méthode de résolution des systèmes linéaires, au point de vue de l'approximation avec laquelle les résultats sont obtenus.

Toute méthode de résolution doit nécessairement conduire à un système d'équations du genre du système VI permettant d'obtenir directement une des inconnues et de déterminer successivement toutes les autres. Supposons qu'on ait été conduit au système :

$$\text{VII}\left\{\begin{array}{l} \delta_1^1\lambda_1 + \delta_1^2\lambda_2 + \cdots + \delta_1^n\lambda_n - \varepsilon_1 = 0 \\ \qquad \delta_2^2\lambda_2 + \cdots + \delta_2^n\lambda_n - \varepsilon_2 = 0 \\ \\ \qquad\qquad \delta_n^n\lambda_n - \varepsilon_n = 0 \end{array}\right.$$

on peut faire correspondre à ce système, un second système fournissant les valeurs des $\varepsilon$, soit :

$$\text{VIII}\left\{\begin{array}{l} \beta_1^1\varepsilon_1 + C_1 = 0 \\ \beta_1^2\varepsilon_1 + \beta_2^2\varepsilon_2 + C_2 = 0 \\ \cdots \\ \beta_1^n\varepsilon_1 + \beta_2^n\varepsilon_2 + \cdots + \beta_n^n\varepsilon_n + C_n = 0 \end{array}\right.$$

les formules IV sont alors remplacées par les suivantes :

$$A_p^p = \sum_{k=1}^{k=p}\left(\beta_k^p\delta_k^p\right)$$

$$A_p^q = \sum_{k=1}^{k=p}\left(\beta_k^p\delta_k^q\right)$$

D'où l'on peut conclure que le système unique des coefficients $\alpha$ que l'on a employé précédemment est remplacé dans tous les autres modes de résolution par un double système de coefficients $\beta$ et $\delta$ tels que l'on a toujours $\beta_p^q\delta_p^q = \left(\alpha_p^q\right)^2$

Or les calculs s'effectuent nécessairement avec une précision limitée et on est amené pour éviter des ~~erreurs~~ fautes et rendre le calcul aussi simple que possible, à calculer tous les nombres employés avec un nombre de décimales fixe. Il en résulte que les nombres $\alpha$, $\beta$, $\delta$ sont affectés d'une erreur $\eta$ dépendant des décimales négligées et indépendante de la grandeur du nombre calculé.

~~L'emploi~~ de la quantité $\left(\alpha_p^q\right)^2$ dans les calculs conduit à l'introduction d'une erreur $2.\alpha_p^q.\eta$

L'emploi de la quantité égale $\left(\beta_p^q\delta_p^q\right)$ correspond à l'introduction de l'erreur $(\beta_p^q + \delta_p^q)\eta$.

On sait que le produit $\beta_p^q \delta_p^q$ étant constant la somme de ses deux facteurs atteint son minimum lorsqu'ils sont égaux. L'erreur la plus faible que l'on puisse introduire est donc

$$2\alpha_p^q \eta.$$

Il en résulte que le mode de résolution des systèmes linéaires qui vient d'être exposé apparaît comme celui qui fournit la meilleure approximation des calculs.

Cette propriété qui permet de réduire au strict minimum le nombre des décimales à employer dans les calculs, compense largement le petit inconvénient d'employer la racine carrée dans la résolution d'équations linéaires.

D'autant plus que la racine carrée peut être obtenue facilement et rapidement à l'aide de la machine à calculer par le procédé suivant complètement différent des procédés généralement indiqués par les fabricants de machines à calculer.

Soit à extraire la racine carrée d'un nombre $N$.

Supposons qu'on connaisse un nombre $n$ voisin de la racine $r$ cherchée.

Soit pour fixer les idées

$$r = n + \varepsilon$$

$$N = r^2 = (n+\varepsilon)^2 = n^2 + 2n\varepsilon + \varepsilon^2$$

Si $\varepsilon^2$ est d'un ordre inférieur à la dernière décimale que l'on veut calculer, on a le droit d'écrire

$$N = n^2 + 2n\varepsilon = n(n+2\varepsilon)$$

c'est à dire qu'en divisant $N$ par $n$ on a pour quotient $n + 2\varepsilon$

$2\varepsilon$ représente l'excès de ce quotient sur le diviseur $n$ et l'on obtient $r$ en ajoutant à $n$ la moitié de cet excès.

Pratiquement, il est avantageux d'avoir à sa disposition une table de carrés qui donne à première vue la racine carrée d'un nombre quelconque avec 3 chiffres significatifs exacts.

$\frac{\varepsilon}{n}$ est alors inférieur à $\frac{1}{10^2}$

$\frac{\varepsilon^2}{n^2}$ —— $\frac{1}{10^4}$

la première division donne la racine avec 5 chiffres significatifs.

$\frac{\varepsilon^4}{n^4}$ est inférieur à $\frac{1}{10^8}$

donc la 2ème division faite avec la racine à 5 chiffres exacts donnerait 9 chiffres significatifs exacts et ainsi de suite.

On peut énoncer une règle simple en supposant que le nombre comporte une virgule placée à la droite du premier chiffre de gauche. Dans ces conditions chaque division double le nombre des décimales de la racine.

A l'aide de ce procédé, un opérateur exercé obtient en quelques secondes la racine carrée d'un nombre de 5 chiffres avec le même nombre de chiffres exacts.

---

La méthode de résolution des systèmes linéaires qui vient d'être exposée a été complétée par l'adaptation du système de vérification indiqué par Gauss sous le nom de preuve par sommes

La vérification est obtenue de la façon suivante : On juxtapose au terme constant $C_p$ de l'équation de rang $p$ un terme $V_p$ donné par la relation $-V_p = A_1^p + A_2^p + \ldots + A_n^p + C_p$

Dans ces conditions la somme des nombres inscrits dans la ligne $p$ du système des équations est nulle. Si l'on traite $V_p$ dans la résolution de la même façon que $C_p$, cette relation linéaire se maintiendra et sera encore vraie pour les coefficients $\alpha$.

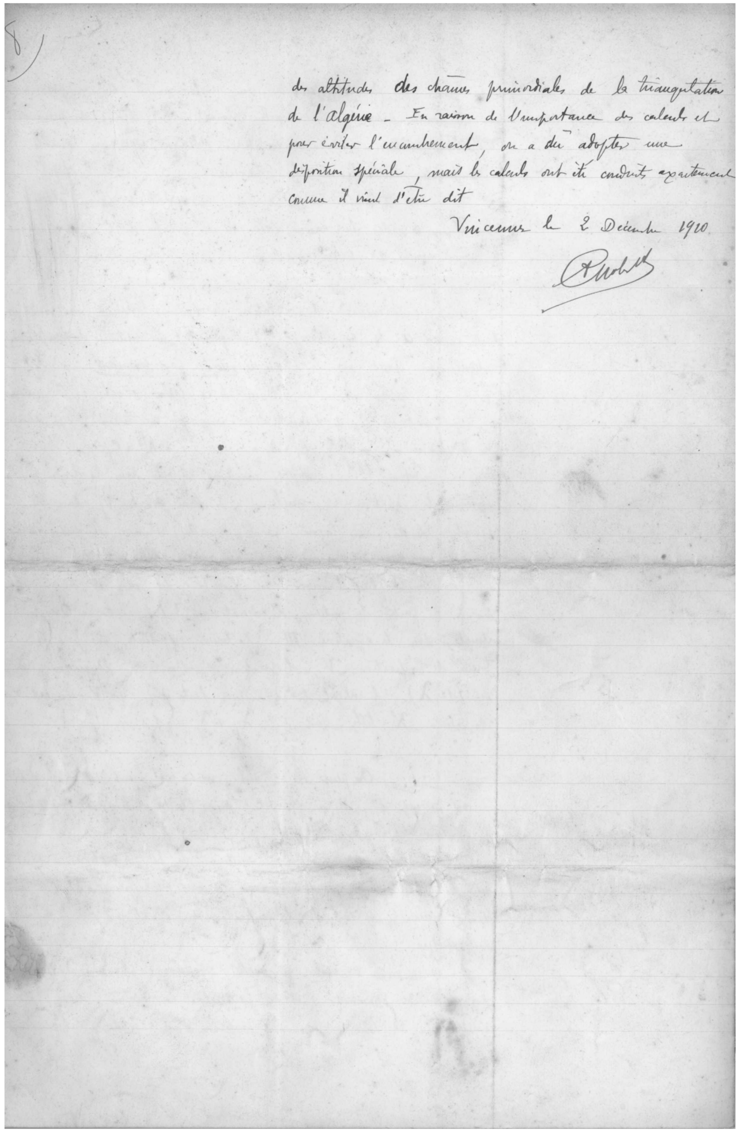
De plus il sera encore possible de vérifier le calcul des $\lambda$ à partir du système des équations VI), car si l'on remplace dans les équations III les termes constants $C$ par les termes de vérification $V$, l'opération équivalant à changer $\lambda$ en $(1-\lambda)$, le calcul des inconnues fait avec les $V$ donnera donc des valeurs $\lambda'$ telles que $\lambda_p + \lambda'_p = 1$

---

On peut en opérant comme il vient d'être dit réussir à coup sûr et en peu de temps la résolution de systèmes d'équations très complexes.

La résolution d'un système de 10 équations à 10 inconnues peut être faite ~~4 à 5 h~~ avec 5 chiffres exacts, en 4 à 5 heures, y compris la vérification des équations et le calcul des résidus.

On a résolu par cette méthode plusieurs systèmes dépassant 30 équations, et en particulier un système de 56 équations. Ce dernier cas fait partie d'un calcul de compensation



8

des altitudes des chaînes primordiales de la triangulation de l'Algérie – En raison de l'importance des calculs et pour éviter l'encombrement, on a dû adopter une disposition spéciale, mais les calculs ont été conduits exactement comme il vient d'être dit

Vincennes le 2 Décembre 1910.

# Translation of the manuscript

## On the numerical solution of systems of linear equations
## A. Cholesky

*The solution of problems depending on experimental data, which, in certain cases, can be subject to conditions, and to which we apply the method of least squares, is always subject to the numerical calculation of the roots of a system of linear equations. This applies to the search of the laws of physics; it is also the case for the compensation of geodetic networks. It is therefore interesting to find a safe and as simple as possible process for the numerical solution of a system of linear equations.*

*The process that we will specify applies to symmetrical systems of equations to which leads the method of least–squares; but we will first notice that the solution of a system of $n$ linear equations in $n$ unknowns can easily be* *reduced to the solution of a system of $n$ symmetric equations in $n$ unknowns.*

*Indeed, let us consider the following system:*

$$I \begin{cases} \alpha_1^1\gamma_1 + \alpha_2^1\gamma_2 + \alpha_3^1\gamma_3 + & \cdots & +\alpha_n^1\gamma_n + C_1 = 0 \\ \alpha_1^2\gamma_1 + \alpha_2^2\gamma_2 + \alpha_3^2\gamma_3 + & \cdots & +\alpha_n^2\gamma_n + C_2 = 0 \\ \dots\dots\dots\dots\dots\dots\dots & \dots & \dots\dots\dots\dots\dots\dots \\ \alpha_1^n\gamma_1 + \alpha_2^n\gamma_2 + & \cdots & +\alpha_n^n\gamma_n + C_n = 0. \end{cases}$$

*Let us perform the linear transformation represented by the system:*

$$
\text{II}\left\{
\begin{array}{l}
\gamma_1 = \alpha_1^1\lambda_1 + \alpha_1^2\lambda_2 + \cdots + \alpha_1^n\lambda_n \\
\gamma_2 = \alpha_2^1\lambda_1 + \alpha_2^2\lambda_2 + \cdots + \alpha_2^n\lambda_n \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\gamma_n = \alpha_n^1\lambda_1 + \alpha_n^2\lambda_2 + \cdots + \alpha_n^n\lambda_n.
\end{array}
\right.
$$

*The system of equations I giving the $n$ unknowns $\gamma$ is replaced by the system III giving the $n$ unknowns $\lambda$ allowing, using II, to compute the values of the $\gamma$'s.*

$$
\text{III}\left\{
\begin{array}{l}
A_1^1\lambda_1 + A_2^1\lambda_2 + \cdots + A_n^1\lambda_n + C_1 = 0 \\
A_1^2\lambda_1 + A_2^2\lambda_2 + \cdots + A_n^2\lambda_n + C_2 = 0 \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
A_1^n\lambda_1 + A_2^n\lambda_2 + \cdots + A_n^n\lambda_n + C_n = 0.
\end{array}
\right.
$$

*Generally, we have*

$$\text{IV)}\quad \begin{aligned} A_p^p &= \sum_{k=1}^{k=n} (\alpha_k^p)^2 \\ A_p^q &= \sum_{k=1}^{k=n} \alpha_k^p \alpha_k^q. \end{aligned}$$

*The coefficient $A_p^q$ is obtained by the product of the coefficients of the rows $p$ and $q$ of the system I which are in the same column and effecting the sum of the products thus obtained in the $n$ columns, which can be symbolically expressed by saying that $A_p^q$ is the product of the row $p$ by the row $q$.*

*As the order of the factors can be reversed in each product,* *we immediately see that*

$$A_p^q = A_q^p$$

*in the determinant of the system III the terms symmetric with respect to the diagonal are equal, i.e. the system of equations with $\lambda$ is symmetric.*

---

*Therefore, we propose to solve a system of equations of the form III. Notice, from what precedes, that the system of equations II, if the unknowns $\gamma$ were assumed to be known, will be a system of equations in the $\lambda$'s, equivalent to the system III. It would therefore be a way to solve the system III if it was possible to find a system I allowing to easily compute the $\gamma$'s.*

*This is what happens if, in the system I*

*the first equation only contains* $\gamma_1$

*the 2nd* $\gamma_1$ *et* $\gamma_2$

*the 3rd* $\gamma_1$, $\gamma_2$ *et* $\gamma_3$

*and so on. We can indeed compute similarly all the* $\gamma$*'s successively from* $\gamma_1$.

*The problem is thus reduced to the search of the system*

$$V)\begin{cases}\alpha_1^1\gamma_1 & +C_1=0\\ \alpha_1^2\gamma_1+\alpha_2^2\gamma_2 & +C_2=0\\ \alpha_1^3\gamma_1+\alpha_2^3\gamma_2+\alpha_3^3\gamma_3 & +C_3=0\\ \dots\dots\dots\dots\dots\dots\dots\dots & \dots\dots\dots\\ \alpha_1^n\gamma_1+\alpha_2^n\gamma_2+\alpha_3^n\gamma_3+\cdots+\alpha_n^n\gamma_n & +C_n=0.\end{cases}$$

*This system being indeed found, the problem becomes very easy since the system II is replaced by the system IV which allows to compute the $\lambda$'s step by step from $\lambda_n$.*

$$\text{VI)} \left\{ \begin{array}{cccccc} \alpha_1^1\lambda_1 + & \alpha_1^2\lambda_2 + & \cdots & \cdots & +\alpha_1^n\lambda_n - \gamma_1 = 0 \\ & \alpha_2^2\lambda_2 + & \alpha_2^3\lambda_3 & \cdots & +\alpha_2^n\lambda_n - \gamma_2 = 0 \\ & & \alpha_3^3\lambda_3 & \cdots & +\alpha_3^n\lambda_n - \gamma_3 = 0 \\ & & \cdots & \cdots & \cdots\cdots\cdots\cdots\cdots \\ & & & & \alpha_n^n\lambda_n - \gamma_n = 0. \end{array} \right.$$

*We will easily compute the coefficients $\alpha$ starting from the coefficients $A$ of the system III, by applying the general relations IV) to the system V. This shows that one can compute row by row all the coefficients of the system VI*

$1^{th}$ *row*

$$\left\{\begin{array}{lll} A_1^1 = (\alpha_1^1)^2 & \textit{d'où} & \alpha_1^1 = \sqrt{A_1^1} \\ A_2^1 = \alpha_1^1\alpha_1^2 & & \alpha_1^2 = \dfrac{A_2^1}{\alpha_1^1} \\ \dots\dots\dots\dots & & \\ A_p^1 = \alpha_1^1\alpha_1^p & & \alpha_1^p = \dfrac{A_p^1}{\alpha_1^1} \\ A_n^1 = \alpha_1^1\alpha_1^n & & \alpha_1^n = \dfrac{A_n^1}{\alpha_1^1} \end{array}\right.$$

$2^{nd}$ *row*

*the* $\alpha_1$ *are already*

*known by the computation*

*of the* $1^{th}$ *row*

$$\left\{\begin{array}{ll} A_2^2 = (\alpha_1^2)^2 + (\alpha_2^2)^2 & \alpha_2^2 = \sqrt{A_2^2 - (\alpha_1^2)^2} \\ A_3^2 = \alpha_1^2\alpha_1^3 + \alpha_2^2\alpha_2^3 & \alpha_2^3 = \dfrac{A_3^2 - \alpha_1^2\alpha_1^3}{\alpha_2^2} \\ A_p^2 = \alpha_1^2\alpha_1^p + \alpha_2^2\alpha_2^p & \alpha_2^p = \dfrac{A_p^2 - \alpha_1^2\alpha_1^p}{\alpha_2^2} \end{array}\right.$$

$p^{th}$ row

All the $\alpha$ with a lower index smaller than $p$ are known from the computation of the preceding rows.

$$\begin{cases} A_p^p = (\alpha_1^p)^2 + (\alpha_2^p)^2 + (\alpha_3^p)^2 + \cdots + (\alpha_{p-1}^p)^2 + (\alpha_p^p)^2 \\ \\ \quad \alpha_p^p = \sqrt{A_p^p - (\alpha_1^p)^2 - (\alpha_2^p)^2 - (\alpha_3^p)^2 - \cdots - (\alpha_{p-1}^p)^2} \\ \\ A_p^q = \alpha_1^p \alpha_1^q + \alpha_2^p \alpha_2^q + \alpha_3^p \alpha_3^q + \cdots + \alpha_{p-1}^p \alpha_{p-1}^q + \alpha_p^p \alpha_p^q \\ \\ \quad \alpha_p^q = \dfrac{A_p^q - \alpha_1^p \alpha_1^q - \alpha_2^p \alpha_2^q - \cdots - \alpha_{p-1}^p \alpha_{q-1}^q}{\alpha_p^p} \\ q > p \end{cases}$$

As for the computation of the $\gamma$'s, it is easily obtained using *the equations V. We obtain*

$$
\begin{array}{rcl}
(-\gamma_1) & = & \dfrac{C_1}{\alpha_1^1} \\
(-\gamma_2) & = & \dfrac{C_2 - \alpha_1^2(-\gamma_1)}{\alpha_2^2} \\
\cdots & \cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
(-\gamma_p) & = & \dfrac{C_p - \alpha_1^p(-\gamma_1) - \alpha_2^p(-\gamma_2) - \cdots}{\alpha_p^p}
\end{array}
$$

*which shows that the coefficients $(-\gamma)$ which appear in the table of the equations VI) are computed with respect to the constant terms $C$ of the table III exactly in the same way as the coefficients $\alpha$ with respect to the $A$'s.*

---

*The computations can be presented in a convenient way* *into a single table. The given equations being symmetric, it is sufficient to write, in the table, the coefficients in a single side of the diagonal, above for example.*

*The transformed equations of system VI can then be displayed under the diagonal symmetrically to the given equations, each new equation occupying a column below the diagonal.*

*The writings are restricted to the replication of the coefficients $\alpha$; indeed, the computation of a coefficient $\alpha$ of the form $\frac{\sum mn}{K}$ is realized on a calculator by a sequence of multiplications which are automatically and algebraically added on the machine, the algebraic sum being immediately divided by $K$. Thus, the most complicated operation that the machine is able to perform is used, then the maximal output of this machine is obtained, while staying as much as possible away from the errors* (deleted and replaced by "faults"), *frequent in the transcription of the digits read.*

*Simplifications in the application of the formulae or precious indications are found in the layout of the computations, and in the use of the calculator as well. Let us give one single example: the calculators of "Dactyle" type record the quotient in white or red digits according if the division is performed by turning* (translator: the crank) *in the direction of addition or substraction, and it follows that every coefficient $\alpha$ being the result of a division on the machine, its sign is indicated by the color of the digits that compose it; sign errors are thus easily avoided.*

*It seems useless to dwell on the rest of the resolution, that is on the computation of the $\lambda$ from the system VI). Indeed, we immediately see how we can turn back from $\lambda_n$ to $\lambda_{n-1}$, then to $\lambda_{n-2}$, and so on until $\lambda_1$.*

*We can highlight the advantages of this method of solving linear systems in terms of the approximation with which the results are obtained.*

*Every solution method must necessarily lead to a system of equations of the type of the system VI allowing to obtain directly one of the unknowns and to determine successively all the others. Assume that we have been led to the system:*

$$\text{VII}\begin{cases} \delta_1^1\lambda_1 + & \delta_1^2\lambda_2 + & \cdots & +\delta_1^n\lambda_n - \varepsilon_1 = 0 \\ & \delta_2^2\lambda_2 + & \cdots & +\delta_2^n\lambda_n - \varepsilon_2 = 0 \\ & & \cdots & \cdots\cdots\cdots\cdots\cdots \\ & & & \delta_n^n\lambda_n - \varepsilon_n = 0 \end{cases}$$

*We can put in correspondence this system with a second* *system giving the values of the $\varepsilon$'s, that is:*

$$\text{VIII}\begin{cases} \beta_1^1\varepsilon_1 & +C_1 = 0 \\ \beta_1^2\varepsilon_1 + \beta_2^2\varepsilon_2 & +C_2 = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots & \dots\dots\dots \\ \beta_1^n\varepsilon_1 + \beta_2^n\varepsilon_2 + \dots + \beta_n^n\varepsilon_n & +C_n = 0. \end{cases}$$

*Formulae IV are then replaced by the following ones:*

$$A_p^p = \sum_{k=1}^{k=p}(\beta_k^p\delta_k^p)$$

$$A_p^q = \sum_{k=1}^{k=p}(\beta_k^p\delta_k^q).$$

*From what we can conclude that the unique system of the coefficients $\alpha$ that we previously used is replaced in all the* *other modes of resolution by a double system of coefficients $\beta$ and $\delta$ such that we always have $\beta_p^q \delta_p^q = (\alpha_p^q)^2$.*

*But the calculations are performed necessarily with a limited accuracy and we are led, to avoid errors* (deleted and replaced by "faults") *and make the computation as simple as possible, to compute all the numbers used with a fixed number of figures. It follows that the numbers $\alpha, \beta, \delta$ are affected by an error $\eta$ depending on the neglected figures and independent of the magnitude of the number calculated.*

*The use of the quantity $(\alpha_p^q)^2$ in the computations corresponds to the introduction of an error $2\alpha_p^q \eta$.*

*The use of the equal quantity $(\beta_p^q \delta_p^q)$ corresponds to the introduction of the error $(\beta_p^q + \delta_p^q)\eta$.*

*We know that the product $\beta_p^q \delta_p^q$ being constant, the sum of its* *two factors attains its minimum when they are equal. Thus, the smallest error that could be introduced is*

$$2\alpha_p^q\eta.$$

*As a result, the mode of resolution of linear systems which has just been exposed appears as the one which provides the best approximation of the computations.*

*This property allows to reduce to its strict minimum the number of figures to be used in the computations, it widely compensates the small inconvenience of using the square root in the solution of linear equations.*

*Especially since the square root can be easily and rapidly obtained with a calculator by the following procedure which is completely different from the processes usually indicated by calculators' manufacturers.*

*Let us extract the square root of a number $N$.*

*Assume that a number $n$ close to the required root $r$ is known.*

*As to fix the ideas*

$$\begin{aligned} r &= n+\varepsilon \\ N &= r^2 = (n+\varepsilon)^2 = n^2 + 2n\varepsilon + \varepsilon^2. \end{aligned}$$

*If $\varepsilon^2$ has its last digit one order less than what we want to compute, we have the right to write*

$$N = n^2 + 2n\varepsilon = n(n+2\varepsilon)$$

*that is, dividing $N$ by $n$, we have as the quotient*

$$n + 2\varepsilon.$$

*$2\varepsilon$ represents the excess of this quotient on the divisor $n$, and*

*we obtain $r$ by adding the half of this excess to $n$.*

*Practically, it is advantageous to have at one's disposal a table of squares which gives at first sight the square root of any number with 3 exact significant figures.*

$\frac{\varepsilon}{n}$ *is then smaller than* $\frac{1}{10^2}$
$\frac{\varepsilon^2}{n^2}$ $\frac{1}{10^4}$.

*The first division gives the square root with 5 significant figures.*

$\frac{\varepsilon^4}{n^4}$ *is smaller than à* $\frac{1}{10^8}$

*thus the $2^{nd}$ division made with the root having 5 exact figures would give 9 exact significant digits and so on.*

*We can state a simple rule assuming that the number has a decimal point at the right of the first digit on the left. Under* *these conditions each division doubles the number of decimal places the root.*

*Using this method, an exercised operator gets in a few seconds the square root of a number of 5 digits with the same number of exact figures.*

---

*The method for solving linear systems which has just been exposed was supplemented by an adaptation of the verification procedure indicated by Gauss under the name of proof by sums.*

*The verification is obtained as follows: we juxtaposes to the constant term $C_p$ of the equation of rank $p$ a term $V_p$ given by the relation $-V_p = A_1^p + A_2^p + \cdots + A_n^p + C_p$.*
*Under these conditions the sum of the numbers listed in the*

*row $p$ of the system of equations is zero. If we treat $V_p$ in the resolution the same way as $C_p$, this linear relation will be preserved and it will still be true for the coefficients $\alpha$.*

*Moreover, it will still be possible to verify the computation of the $\lambda$'s from the system of equations VI), since if we replace in the equations III the constant terms $C$ by the checking terms $V$, the operation is equivalent to changing $\lambda$ into $(1-\lambda)$; the computation of the unknowns made with the $V$'s therefore gives values $\lambda'$ such that $\lambda_p + \lambda'_p = 1$.*

---

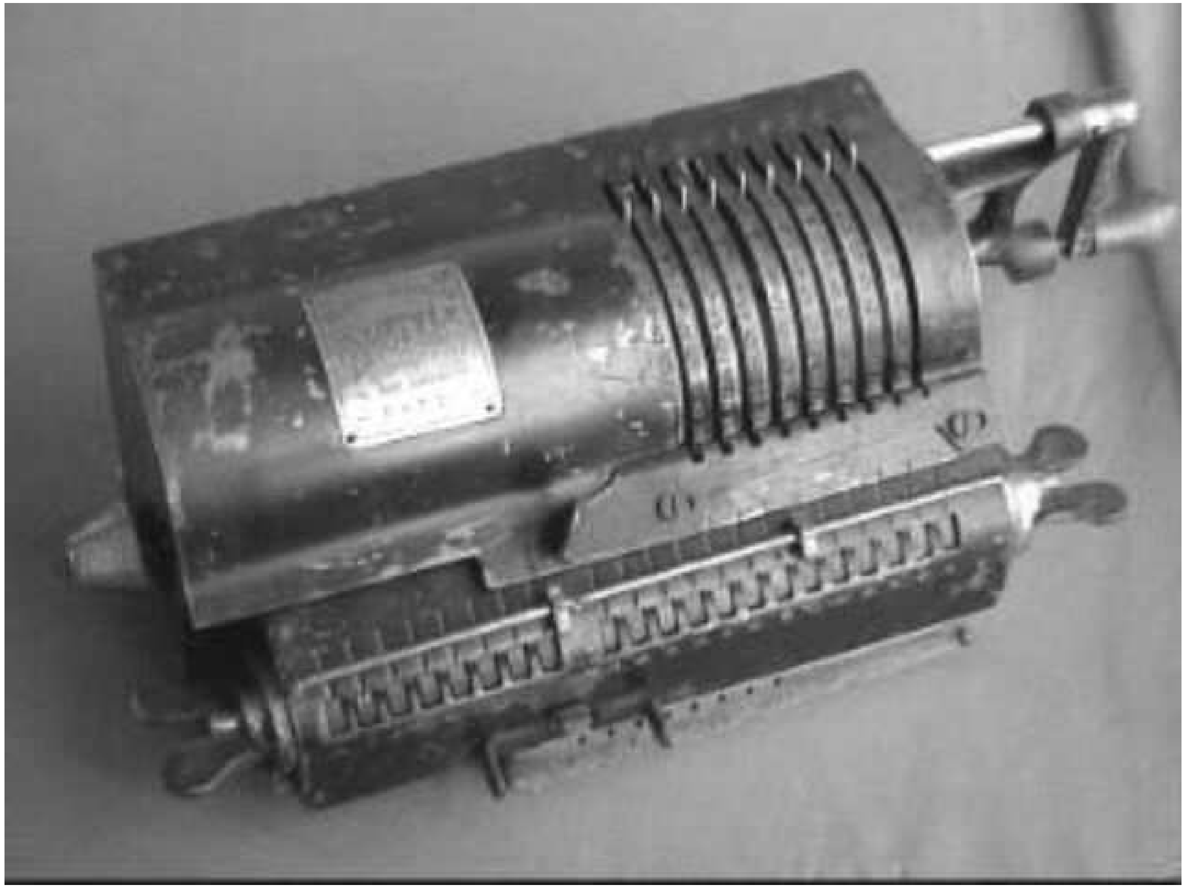
*It is possible, operating as just said, to succeed for sure and in a short time the resolution of very complex systems of equations.*

*The solution of a system of 10 equations with 10 unknowns* *can be obtained with 5 exact figures in 4 to 5 hours, including the verification of the equations and the computation of the residuals.*

*By this method, several systems over 30 equations were solved and, in particular, a system of 56 equations. This last case is part of a compensation calculation of the altitudes of the primordial chains of the triangulation of Algeria. Because of the importance of calculations and to avoid congestion, we had to adopt a special arrangement, but the computations have been conducted exactly as has been said.*

*Vincennes 2 December 1910*

*Signature*

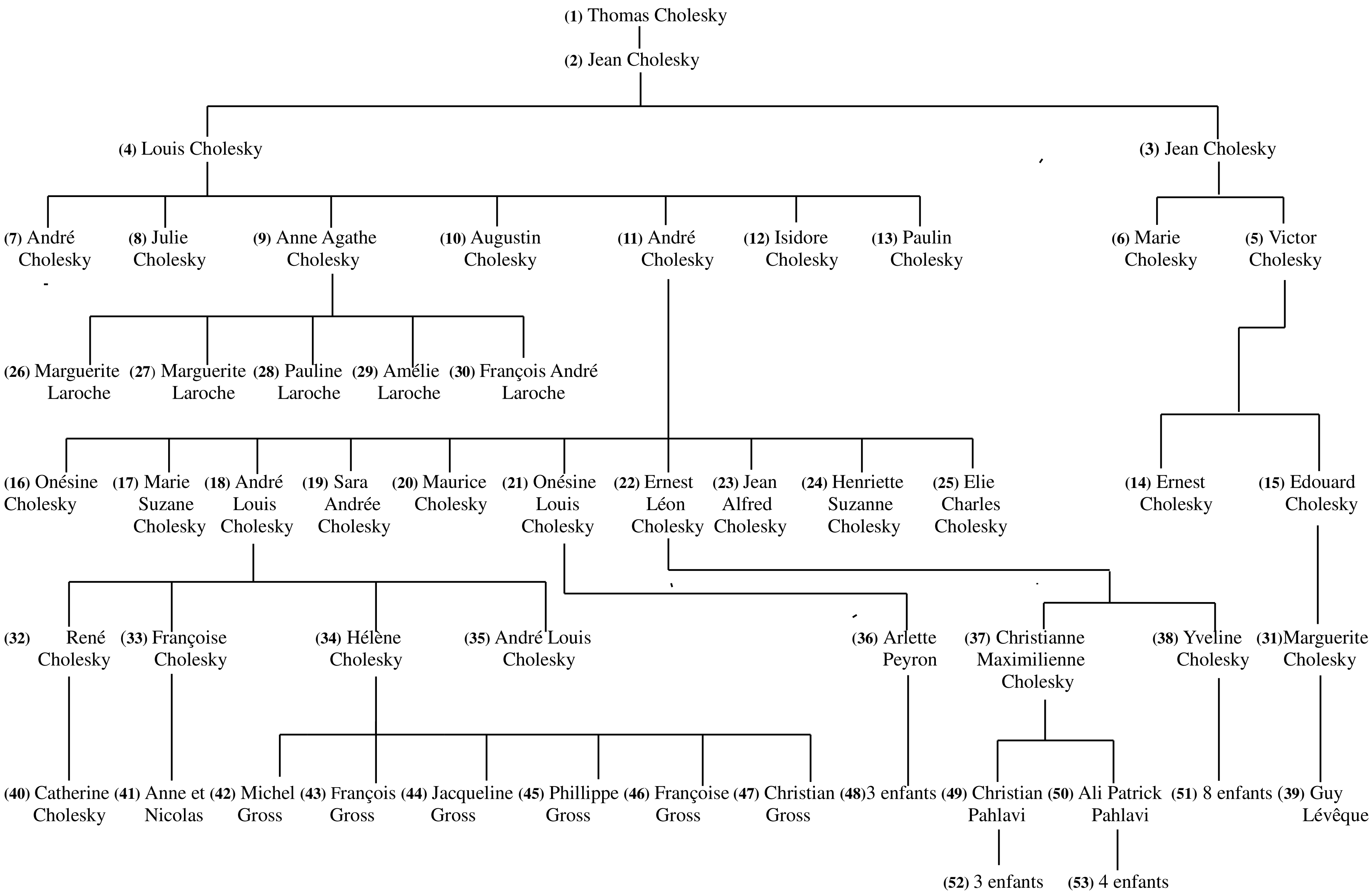


## The family

The first member of the family encountered in the archives is **Thomas Cholesqui**, with this spelling, around 1770.

It seemed to be issued from a noble Polish family. But this is not sure since some documents mention Pozsoni (Presbourg in French, and then Bratislava after 1919), the capital of Hungary during the 18th century.

According to some other sites, the family could come from Ukraine.


(1) Thomas Cholesky
(2) Jean Cholesky
(4) Louis Cholesky
(3) Jean Cholesky
(7) André Cholesky
(8) Julie Cholesky
(9) Anne Agathe Cholesky
(10) Augustin Cholesky
(11) André Cholesky
(12) Isidore Cholesky
(13) Paulin Cholesky
(6) Marie Cholesky
(5) Victor Cholesky
(26) Marguerite Laroche
(27) Marguerite Laroche
(28) Pauline Laroche
(29) Amélie Laroche
(30) François André Laroche
(16) Onésine Cholesky
(17) Marie Suzane Cholesky
(18) André Louis Cholesky
(19) Sara Andrée Cholesky
(20) Maurice Cholesky
(21) Onésine Louis Cholesky
(22) Ernest Léon Cholesky
(23) Jean Alfred Cholesky
(24) Henriette Suzanne Cholesky
(25) Elie Charles Cholesky
(14) Ernest Cholesky
(15) Edouard Cholesky
(32) René Cholesky
(33) Françoise Cholesky
(34) Hélène Cholesky
(35) André Louis Cholesky
(36) Arlette Peyron
(37) Christianne Maximilienne Cholesky
(38) Yveline Cholesky
(31)Marguerite Cholesky
(40) Catherine Cholesky
(41) Anne et Nicolas
(42) Michel Gross
(43) François Gross
(44) Jacqueline Gross
(45) Phillippe Gross
(46) Françoise Gross
(47) Christian Gross
(48)3 enfants
(49) Christian Pahlavi
(50) Ali Patrick Pahlavi
(51) 8 enfants
(39) Guy Lévêque
(52) 3 enfants
(53) 4 enfants