

Werk

Label: Article

Jahr: 1939

PURL: https://resolver.sub.uni-goettingen.de/purl?31311028X_0068|log57

Kontakt/Contact

[Digizeitschriften e.V.](#)
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

✉ info@digizeitschriften.de

Sur une méthode approximative de compensation des fonctions empiriques.

Zdeněk Horák, Praha.

(Reçu le 19 décembre 1938.)

La solution du problème de la compensation des fonctions empiriques par la méthode des moindres carrés présente l'inconvénient d'un calcul numérique laborieux. C'est pour cela que l'on a imaginé plusieurs méthodes approximatives plus faciles à utiliser. Dans le présent mémoire, je me propose de traiter une nouvelle méthode approximative du calcul des compensations des fonctions paramétriques que j'ai déjà appliquée aux résultats de mes mesures relatives à la conductibilité thermique et à la viscosité en fonction de la température. La méthode se révéla commode et suffisamment précise.

1. Principe de la méthode. Considérons une quantité x fonction de la variable t

$$x = f(t, a_1, a_2, \dots, a_p) \quad (1)$$

dans laquelle a_1, a_2, \dots, a_p sont des paramètres de valeurs inconnues. Pour les valeurs t_1, t_2, \dots, t_m de t nous mesurons les valeurs x_1, x_2, \dots, x_m de x . Le problème consiste à déterminer les meilleures valeurs à adopter pour les paramètres a_1, \dots, a_p , en tenant compte de résultats de mesure fournissant un nombre surabondant de relations entre ces paramètres. Or, la méthode des moindres carrés donne pour les paramètres a_1, \dots, a_p les équations

$$\frac{\partial}{\partial a_r} \sum_{i=1}^m [x_i - f(t_i, a_1, \dots, a_p)]^2 = 0, \quad r = 1, 2, \dots, p \quad (2)$$

dont l'établissement et la résolution sont, en général, assez pénibles. Cependant, lorsqu'un des paramètres, a_1 par exemple, intervient dans la fonction f comme une constante additive, la première des équations (2) — pour $r = 1$ — prend la forme relativement très simple

$$\sum_{i=1}^m [x_i - f(t_i, a_1, \dots, a_p)] = 0. \quad (3)$$

On peut profiter de ce fait, en procédant comme suit:

Divisons toutes les mesures en p groupes et remplaçons les p équations (2) par celles (3) appliquées aux p groupes de mesures:

$$\sum_{(1)} (x - f) = 0, \sum_{(2)} (x - f) = 0, \dots, \sum_{(p)} (x - f) = 0 \quad (4)$$

où la somme $\sum_{(r)}$ concerne le r -ième groupe. Si les mesures effectuées satisfont approximativement à la relation (1), on peut s'attendre à ce que les valeurs des paramètres a_1, \dots, a_p , obtenues en résolvant les équations (4), ne diffèrent pas beaucoup de celles calculées d'après les équations (2). Néanmoins, il faut admettre que les paramètres a_1, \dots, a_p définis par (4) dépendent du nombre de mesures contenues dans chaque groupe (*répartition quantitative*) et de la manière dont on répartit les mesures individuelles dans les groupes (*répartition qualitative*). Il importe donc de trouver les répartitions quantitative et qualitative des mesures les plus avantageuses, fournissant les valeurs des paramètres les plus exactes, ce dont je parlerai dans la suite.

2. Condition fondamentale pour le choix des groupes de mesures. Nous partirons de la supposition, adoptée ci-dessus, que le paramètre a_1 intervient dans la fonction f comme une constante additive c'est-à-dire que l'équation (1) est de la forme

$$x = a_1 + \varphi(t, a_2, \dots, a_p) \quad (5)$$

où φ est une fonction continue et uniforme d'ailleurs quelconque. Donc les équations (4) prennent la forme suivante

$$\left. \begin{aligned} \sum_{(1)} x_i &= m_{(1)} a_1 + \sum_{(1)} \varphi(t_i, a_2, \dots, a_p), \\ \sum_{(2)} x_i &= m_{(2)} a_1 + \sum_{(2)} \varphi(t_i, a_2, \dots, a_p), \\ &\dots\dots\dots \\ \sum_{(p)} x_i &= m_{(p)} a_1 + \sum_{(p)} \varphi(t_i, a_2, \dots, a_p), \end{aligned} \right\} \quad (6)$$

$m_{(1)}, m_{(2)}, \dots, m_{(p)}$ désignant les nombres de mesures contenues dans le premier, le second, ... le r -ième groupe, de sorte que

$$m_{(1)} + m_{(2)} + \dots + m_{(p)} = m. \quad (7)$$

Si l'on porte les valeurs t_i, x_i dans l'équation hypothétique (5), on arrive aux équations

$$x_i = a_1 + \varphi(t_i, a_2, \dots, a_p), \quad i = 1, \dots, m \quad (5')$$

auxquelles elles ne satisfont qu'approximativement. Cependant

leurs x_i . Par suite, il y a intérêt à avoir les seconds membres possible-ment précis sans s'inquiéter des erreurs de détermination de pre-miers membres des relations (12). Je me propose donc de chercher un tel choix de groupes de mesures qui donne la plus grande préci-sion relative des seconds membres des équations (12) au moyen desquelles nous calculons les paramètres inconnus a_2, \dots, a_p . Autrement dit, *je vais considérer comme le plus avantageux le choix des groupes de mesures pour lequel les erreurs relatives des seconds membres des équations (12) sont minima.*

Désignons par $\delta_{(r)}$ l'erreur quadratique moyenne des mesures du r -ième groupe. Alors l'erreur quadratique moyenne $\bar{\delta}_{(r)}$ de la moyenne arithmétique $\bar{x}_{(r)}$ est donnée, comme nous le savons, par l'expression

$$\bar{\delta}_{(r)} = \frac{\delta_{(r)}}{\sqrt{m_{(r)}}}$$

et celle de la différence $\bar{x}_{(r)} - \bar{x}_{(s)}$ a pour valeur

$$\sqrt{\bar{\delta}_{(r)}^2 + \bar{\delta}_{(s)}^2} = \sqrt{\frac{\delta_{(r)}^2}{m_{(r)}} + \frac{\delta_{(s)}^2}{m_{(s)}}}. \quad (13)$$

En divisant par la différence elle-même, on obtient l'erreur relative.

On a donc la relation à satisfaire

$$\frac{1}{\bar{x}_{(r)} - \bar{x}_{(s)}} \sqrt{\frac{\delta_{(r)}^2}{m_{(r)}} + \frac{\delta_{(s)}^2}{m_{(s)}}} = \text{minimum} \quad (14)$$

remplaçable par les deux conditions simultanées:

$$\frac{\delta_{(r)}^2}{m_{(r)}} + \frac{\delta_{(s)}^2}{m_{(s)}} = \text{minimum}, \quad (14a)$$

$$\bar{x}_{(r)} - \bar{x}_{(s)} = \text{maximum}, \quad (14b)$$

pouvant être regardées comme indépendantes. En effet, on peut évidemment supposer que toutes les mesures sont à peu près de la même précision et les erreurs $\delta_{(r)}$ ont donc — en première appro-ximation — la même valeur pour tous les groupes. Donc l'erreur $\bar{\delta}_{(r)}$ de la moyenne arithmétique $\bar{x}_{(r)}$ dépend très peu du choix des mesures contenues dans le r -ième groupe, mais elle est une fonction du nombre $m_{(r)}$ de ces mesures. Alors, l'expression (14a), qui dépend explicitement des nombres $m_{(r)}, m_{(s)}$, est presque indépendante de la répartition qualitative des mesures dans les deux groupes corres-pondants. D'autre part, la valeur de la moyenne arithmétique $\bar{x}_{(r)}$ dépend essentiellement de cette répartition, c'est-à-dire du choix des mesures individuelles contenues dans le r -ième groupe, tandis que le nombre $m_{(r)}$ peut varier dans une large mesure sans avoir

grande influence sur la valeur de $\bar{x}_{(r)}$. Cela permet de considérer l'expression (14b) comme pratiquement indépendante des nombres $m_{(r)}$, $m_{(s)}$. Il en résulte en somme que les deux conditions (14a) et (14b) sont équivalentes à la condition fondamentale (14) et peuvent être satisfaites, indépendamment l'une de l'autre de la manière suivante: On détermine les nombres $m_{(r)}$, $m_{(s)}$ de sorte que la condition (14a) soit remplie sans tenir compte de la condition (14b) et réciproquement on détermine la répartition qualitative des mesures dans les deux groupes qui satisfasse à la condition (14b), les nombres $m_{(r)}$, $m_{(s)}$, supposés constants.

3. Répartition quantitative des mesures. Nous commençons par la condition (14a), exprimant que l'erreur quadratique moyenne, donnée par (13), de la différence $\bar{x}_{(r)} - \bar{x}_{(s)}$ doit être minima. Cette condition conduit à $p - 1$ équations indépendantes qui s'obtiennent pour les différentes valeurs des indices r , s entre 1 et p . Par raison de symétrie et d'indépendance de ces équations, il suffit évidemment d'en résoudre une seule en supposant que les nombres de mesures obtenues dans les autres groupes, qui n'interviennent pas dans l'équation en question, restent invariables. Si nous choisissons l'équation (14a) elle-même, les nombres $m_{(k)}$ peuvent être regardés comme constants pour toutes les valeurs de k sauf r et s . Donc, en vertu de (7),

$$m_{(r)} + m_{(s)} = C^{te}. \quad (15)$$

En annulant la différentielle du premier membre de (14a), on aura

$$\frac{\delta_{(r)}^2}{m_{(r)}^2} dm_{(r)} + \frac{\delta_{(s)}^2}{m_{(s)}^2} dm_{(s)} = 0$$

et d'après (15)

$$dm_{(r)} + dm_{(s)} = 0$$

ce qui donne

$$\frac{\delta_{(r)}^2}{m_{(r)}^2} - \frac{\delta_{(s)}^2}{m_{(s)}^2} = 0.$$

Comme les $m_{(r)}$ et $m_{(s)}$ sont essentiellement positifs, on trouve enfin

$$m_{(r)} : m_{(s)} = \delta_{(r)} : \delta_{(s)}. \quad (16)$$

Dans cette relation, les indices r , s peuvent être choisis arbitrairement, ce qui entraîne

$$m_{(1)} : m_{(2)} : \dots : m_{(p)} = \delta_{(1)} : \delta_{(2)} : \dots : \delta_{(p)}. \quad (16')$$

Cette dernière relation et l'équation (7) déterminent complètement les nombres $m_{(1)}$, \dots , $m_{(p)}$, si l'on connaît les rapports des erreurs quadratiques moyennes $\delta_{(k)}$. Or, j'ai déjà signalé que les erreurs $\delta_{(k)}$ sont à peu près égales et c'est en le supposant que j'ai déduit la relation (16'). Donc, les erreurs quadratiques moyennes des seconds

membres des équations (12), définissant les paramètres a_2, \dots, a_p , sont minima, lorsqu'on choisit tous les nombres $m_{(k)}$ approximativement égaux:

$$m_{(1)} \doteq m_{(2)} \doteq \dots \doteq m_{(p)} \doteq \frac{m}{p}. \quad (17)$$

Le nombre m des mesures n'est pas divisible, en général, par le nombre p des paramètres, donc les nombres $m_{(k)}$ ne peuvent être rigoureusement égaux si nous exigeons qu'ils soient entiers. Dans le cas où le nombre des mesures est assez grand par rapport au nombre des paramètres, on peut satisfaire aux équations (17) avec une approximation suffisante. Autrement il faut distinguer deux cas particuliers:

1° *Les mesures sont d'égale précision.* Alors, on choisit tous les nombres $m_{(k)}$ égaux au rapport m/p , somme d'un nombre entier e et d'une fraction u/v , c'est-à-dire qu'on forme les p équations (6), en faisant p sommes à e équations (5'), en ajoutant à chaque somme une des équations (5') superflues multipliée par u/v ou plusieurs d'entre elles multipliées par des coefficients dont la somme est égale à la fraction u/v .

Si par exemple $m = 20$ et $p = 3$, on a $m/p = 6\frac{2}{3}$ et l'on peut choisir les sommes suivantes:

$$1^{\text{er}} \text{ groupe: } E_1 + E_2 + \dots + E_6 + \frac{2}{3}E_7,$$

$$2^{\text{e}} \text{ groupe: } \frac{1}{3}E_7 + E_8 + \dots + E_{13} + \frac{1}{3}E_{14},$$

$$3^{\text{e}} \text{ groupe: } \frac{2}{3}E_{14} + E_{15} + \dots + E_{19} + E_{20},$$

où le symbole E_l signifie la l -ième équation du système (5').

2° *Les mesures sont d'inégale précision.* Dans ce cas-là, il est inutile d'introduire des nombres fractionnaires, car il est plus avantageux, comme d'ailleurs l'exige la proportion (16'), de choisir les nombres $m_{(k)}$ plus grands pour les groupes contenant des mesures moins précises. Supposons, pour fixer les idées, que la précision s'abaisse avec l'indice croissant de la mesure. Soit, par exemple, $m = 40$, $p = 3$. Alors, on peut prendre: $m_{(1)} = 12$, $m_{(2)} = 13$, $m_{(3)} = 15$ et former les équations (6) d'après le schéma suivant:

$$1^{\text{er}} \text{ groupe: } E_1 + E_2 + \dots + E_{12},$$

$$2^{\text{e}} \text{ groupe: } E_{13} + E_{14} + \dots + E_{25},$$

$$3^{\text{e}} \text{ groupe: } E_{26} + E_{27} + \dots + E_{40}.$$

Les erreurs des mesures ne peuvent être déterminées exactement que la compensation une fois faite. Il faut donc juger de la précision des mesures d'après les conditions de mesure ou d'après les positions relatives des points représentant graphiquement la fonction cherchée. Si l'on connaît les poids des mesures x_i , on peut en tenir compte, lorsqu'on forme les équations (6) en formant les

sommes des équations (5') multipliées par les poids des mesures correspondantes. En même temps, il faut, bien entendu, modifier les nombres m et $m_{(k)}$ conformément aux poids adoptés.

4. Répartition qualitative des mesures. Il nous reste encore à satisfaire à la seconde condition (14b). Comme nous l'avons déjà vu, les valeurs $\bar{x}_{(r)}$ sont pratiquement indépendantes des nombres $m_{(r)}$. Ceci est vrai a fortiori pour leurs différences et il y a même des cas où l'on constate qu'elles sont rigoureusement indépendantes des changements des nombres $m_{(k)}$.

Considérons, par exemple, le cas où les valeurs des x_i contenus dans les deux premiers groupes, forment une progression arithmétique:

$$x_i = x_1 + (i - 1) d.$$

Pour la répartition

1^{er} groupe: $x_1, x_2, \dots, x_{m_{(1)}}$,

2^e groupe: $x_{m_{(1)}+1}, \dots, x_{m_{(1)}+m_{(2)}}$,

on déduit aisément

$$\bar{x}_{(1)} = x_1 + (m_{(1)} - 1) \frac{d}{2},$$

$$\bar{x}_{(2)} = x_1 + m_{(1)} d + (m_{(2)} - 1) \frac{d}{2}.$$

Donc la différence

$$\bar{x}_{(2)} - \bar{x}_{(1)} = (m_{(1)} + m_{(2)}) \frac{d}{2}$$

ne dépend que de la somme des nombres de mesures. Pour la répartition choisie, le changement des nombres $m_{(1)}, m_{(2)}$, leur somme supposée constante, laisse la différence $\bar{x}_{(2)} - \bar{x}_{(1)}$ invariable.

Pour remplir la condition (14b), il suffit donc de chercher, pour les nombres $m_{(k)}$ donnés, une répartition qualitative des mesures dans les groupes qui rende maxima les différences $\bar{x}_{(r)} - \bar{x}_{(s)}$. Il est difficile de donner une méthode détaillée et tout-à-fait générale, pour ranger les mesures par groupes, qui nous garantisserait automatiquement — pour toutes les fonctions imaginables — que chacune des différences $\bar{x}_{(r)} - \bar{x}_{(s)}$ a vraiment la plus grande valeur possible et qu'il n'existe pas, dans certains cas particuliers, d'autres répartitions également ou peut-être même plus avantageuses. Je vais déduire cependant une règle très générale, définissant une répartition des mesures qui peut être regardée comme la plus avantageuse et cela pour un vaste champ d'applications.

Considérons la différence $\bar{x}_{(r)} - \bar{x}_{(s)}$ dépendant uniquement de la répartition des mesures dans les deux groupes correspondants

(r -ième et s -ième) et supposons que la disposition des autres mesures par groupes restants soit fixée. Si l'on admet de plus que les nombres $m_{(r)}$, $m_{(s)}$ restent invariables, on s'assure aisément que la valeur absolue de la différence $\bar{x}_{(r)} - \bar{x}_{(s)}$ est maxima, lorsqu'on répartit les mesures de manière que tous les x_i du r -ième groupe soient supérieurs à ceux du s -ième groupe (ou réciproquement). En effet, si l'on substitue l'une à l'autre deux mesures, n'appartenant pas au même groupe, la moyenne $\bar{x}_{(r)}$ ($> \bar{x}_{(s)}$) diminue, tandis que l'autre $\bar{x}_{(s)}$ augmente et donc leur différence subit une diminution. Le même raisonnement peut être fait pour les autres différences et donc, pour un numérotage convenable des groupes, la répartition, satisfaisant le mieux possible à la condition (14b), est caractérisée par le fait que chaque valeur x_i appartenant au k -ième groupe est plus grande que les valeurs de toutes les mesures appartenant au $(k - 1)$ -ième groupe et ainsi de suite.

On peut d'ailleurs remarquer que dans les cas concrets, il n'est pas difficile (en se servant de la représentation graphique) de modifier éventuellement la répartition indiquée pour augmenter, si c'est possible, les différences $\bar{x}_{(r)} - \bar{x}_{(s)}$.

5. Résultat général et application pratique. On voit, en somme, que la répartition qualitative des mesures est déterminée par la condition (14b), tandis que la répartition quantitative est donnée par la relation (17), découlant de la condition (14a). Par suite, la condition fondamentale (14), réunissant les deux conditions précédentes et exprimant que la précision relative des seconds membres des équations (12) doit être maxima, détermine quantitativement et qualitativement la répartition des mesures par groupes. Tenant compte des résultats obtenus, on peut énoncer la règle générale suivante:

On range toutes les m mesures selon la grandeur croissante des valeurs x_1, x_2, \dots, x_m (correspondant aux valeurs t_1, t_2, \dots, t_m de t) et on en forme p groupes également nombreux de manière que chacun d'eux contienne m/p mesures consécutives. Lorsqu'on suppose l'équation, exprimant la quantité x en fonction de la variable indépendante t sous la forme

$$x = a_1 + \varphi(t, a_2, \dots, a_p), \quad (5)$$

les valeurs les plus avantageuses des paramètres a_1, a_2, \dots, a_p sont déterminées par les p relations

$$\bar{x}_{(k)} = a_1 + \bar{\varphi}_{(k)}(t, a_2, \dots, a_p), \quad k = 1, \dots, p, \quad (11)$$

où les $\bar{x}_{(k)}$, $\bar{\varphi}_{(k)}$ désignent les moyennes arithmétiques formées pour le k -ième groupe.

Pour se rendre compte de la précision des résultats obtenus, on procède comme suit: On porte les valeurs des paramètres

a_1, \dots, a_p , tirées de (11) et les valeurs t_1, \dots, t_m dans l'équation (5) ce qui donne les valeurs compensées x'_1, \dots, x'_m de x :

$$x'_i = a_1 + \varphi(t_i, a_2, \dots, a_p). \quad (17)$$

On calcule alors les résidus

$$\varepsilon_i = x'_i - x_i$$

qui satisfont aux relations

$$\sum_{(1)} \varepsilon_i = 0, \sum_{(2)} \varepsilon_i = 0, \dots, \sum_{(p)} \varepsilon_i = 0, \sum_{i=1}^m \varepsilon_i = 0, \quad (18)$$

car, en vertu de (17), (6), (8)

$$\begin{aligned} \sum_{(r)} x'_i &= m_{(r)} a_1 + \sum_{(r)} \varphi(t_i, a_2, \dots, a_p) = \sum_{(r)} x_i, \quad r = 1, \dots, p, \\ \sum_{i=1}^m x'_i &= m a_1 + \sum_{i=1}^m \varphi(t_i, a_2, \dots, a_p) = \sum_{i=1}^m x_i. \end{aligned}$$

D'après les définitions connues, on obtient donc l'erreur quadratique moyenne des mesures du r -ième groupe

$$\delta_{(r)} = \sqrt{\frac{\sum_{(r)} \varepsilon_i^2}{m_{(r)} - 1}} = \sqrt{\frac{[\varepsilon\varepsilon]_{(r)}}{m_{(r)} - 1}} \quad (19)$$

et celle commune à toutes les mesures

$$\delta = \sqrt{\frac{\sum_{i=1}^m \varepsilon_i^2}{m - p}} = \sqrt{\frac{[\varepsilon\varepsilon]}{m - p}}. \quad (20)$$

Connaissant les erreurs $\delta_{(r)}$ nous pouvons vérifier la supposition $\delta_{(r)} \doteq \delta_{(s)}$ et modifier, en cas de besoin, la répartition des mesures par groupes. L'erreur commune δ sert à déterminer les erreurs quadratiques moyennes relatives aux paramètres a_k qu'on obtiendra par les méthodes connues du calcul des compensations.

Quant à l'application de notre méthode, il faut avouer que nous l'avons établie en supposant au no 1, que la fonction f , exprimant la dépendance de x de la variable indépendante t , est de la forme (5), c'est-à-dire qu'un des paramètres inconnus intervient dans f comme constante additive. Cela ne signifie pas nécessairement que le domaine d'application de notre méthode soit en réalité restreint à des fonctions du type (5), il est plutôt plausible d'admettre que la dite supposition n'est pas essentielle pour notre méthode et que les résultats obtenus demeurent valables encore pour n'importe quelles fonctions. Or, cette question n'a guère d'importance pratique, car la supposition adoptée admet l'appli-

cation de la méthode aux fonctions algébriques entières. *Notre méthode est donc directement applicable à toutes les fonctions que l'on peut développer en série des puissances de la variable indépendante.* C'est le cas de l'interpolation parabolique, habituel dans la pratique; en remplaçant la fonction par un polynome entier, on facilite le calcul des paramètres inconnus. C'est pourquoi on préfère, dans le calcul des compensations, les fonctions qui sont linéaires par rapport aux paramètres à déterminer et on se sert souvent du développement en série.

Dans certains cas, on peut donner à la fonction f la forme (5) au moyen d'une transformation convenable. Par exemple, si l'on prend le logarithme d'une fonction exponentielle générale

$$f = ab^{u(t)},$$

on obtient une fonction du type (5), linéaire par rapport aux paramètres $\log a$, $\log b$.

6. Cas des fonctions algébriques entières. — Fonction linéaire.

Au numéro précédent, nous avons constaté que le cas le plus général est pratiquement celui des fonctions algébriques entières. En appliquant la règle générale que j'y ai établie on procèdera donc comme suit:

Dans le cas d'une fonction algébrique entière du degré n -ième

$$x = A_0 + A_1 t + A_2 t^2 + \dots + A_n t^n, \quad (21)$$

on range les mesures d'après la grandeur croissante des valeurs x_i et on en forme $n + 1$ groupes dont le premier contient les premières $\frac{m}{n + 1}$ mesures, le deuxième les $\frac{m}{n + 1}$ mesures suivantes et ainsi de suite. Alors les valeurs les plus avantageuses des paramètres découlent des $n + 1$ équations

$$\bar{x}_{(r)} = A_0 + A_1 \bar{t}_{(r)} + A_2 \bar{t}_{(r)}^2 + \dots + A_n \bar{t}_{(r)}^n, \quad r = 1, \dots, n + 1 \quad (22)$$

où l'on a posé

$$\bar{t}_{(r)}^q = \frac{1}{m_{(r)}} \sum_{(r)} t_i^q = \frac{1}{m_{(r)}} [t^q]_{(r)}.$$

Ces équations sont linéaires par rapport aux inconnus et leur solution ne présente donc pas de difficultés. Il est recommandé d'éliminer A_0 en formant n différences du type

$$(\bar{t}_{(r)} - \bar{t}_{(s)}) A_1 + \dots + (\bar{t}_{(r)}^n - \bar{t}_{(s)}^n) A_n = \bar{x}_{(r)} - \bar{x}_{(s)}. \quad (23)$$

Après en avoir tiré les n paramètres A_1, \dots, A_n , on peut calculer A_0 de l'une quelconque des équations (22) ou bien de l'équation

$$(n + 1) A_0 = \sum_{r=1}^{n+1} \bar{x}_{(r)} - A_1 \sum_{r=1}^{n+1} \bar{t}_{(r)} - \dots - A_n \sum_{r=1}^{n+1} \bar{t}_{(r)}^n \quad (24)$$

$$\left. \begin{aligned} x_I &= A + Bt_I, \\ x_{II} &= A + Bt_{II}. \end{aligned} \right\} \quad (27)$$

Conformément à la règle générale, donnée au no 5, on effectue l'interpolation linéaire de la manière suivante:

On range toutes les mesures par ordre de grandeurs croissantes et l'on calcule les moyennes arithmétiques t_I et x_I des t_i et des x_i correspondant à la première moitié des mesures. Après avoir fait de même pour la seconde moitié, on obtient les moyennes t_{II} , x_{II} . Alors on a deux couples de valeurs de t et de x qui déterminent les deux paramètres de la fonction linéaire cherchée au moyen des équations (27).

En résolvant ces équations, on a

$$B = \frac{x_{II} - x_I}{t_{II} - t_I}, \quad A = \frac{x_I + x_{II}}{2} - B \frac{t_I + t_{II}}{2}, \quad (28)$$

ou aussi

$$A = \bar{x} - \bar{B}t, \quad (29)$$

et par suite les erreurs quadratiques moyennes des paramètres s'écrivent:

$$\delta_B = \frac{1}{t_{II} - t_I} \sqrt{\frac{\delta_I^2}{m_I} + \frac{\delta_{II}^2}{m_{II}}}, \quad \delta_A = \sqrt{(\delta_{\bar{x}})^2 + (\bar{t} \delta_B)^2}, \quad (30)$$

si l'on suppose, bien entendu, que les erreurs des valeurs t_i de la variable indépendante sont négligeables devant celles des valeurs x_i . Les erreurs δ_I , δ_{II} sont données par l'expression (19)

$$\delta_I = \sqrt{\frac{[\varepsilon\varepsilon]_I}{m_I - 1}}, \quad \delta_{II} = \sqrt{\frac{[\varepsilon\varepsilon]_{II}}{m_{II} - 1}} \quad (31)$$

et

$$\delta_{\bar{x}} = \sqrt{\frac{[\varepsilon\varepsilon]}{m(m-2)}}. \quad (32)$$

Dans le cas, d'ailleurs très fréquent, où $m_I = m_{II} = \frac{m}{2}$, on trouve

$$\frac{\delta_I^2}{m_I} + \frac{\delta_{II}^2}{m_{II}} = \frac{4[\varepsilon\varepsilon]}{m(m-2)}$$

ce qui entraîne en vertu de (30)

$$\delta_B = \frac{2}{t_{II} - t_I} \sqrt{\frac{[\varepsilon\varepsilon]}{m(m-2)}} = \frac{2\delta_{\bar{x}}}{t_{II} - t_I}, \quad (33)$$

$$\delta_A = \delta_B \sqrt{\frac{(t_{II} - t_I)^2}{4} + \bar{t}^2} = \delta_B \sqrt{\frac{1}{2}(t_I^2 + t_{II}^2)}. \quad (34)$$

La méthode que je viens d'exposer peut être facilement transformée en une méthode purement graphique, si l'on repré-

sente les mesures dans un système de coordonnées rectangulaires. Cela fait, les points de coordonnées

$$I(t_I, x_I), II(t_{II}, x_{II})$$

sont les centres de gravité respectivement de la première et de la seconde moitié des points représentant les mesures effectuées. Or, le centre de gravité d'un groupe des points est facile à construire, ce qui conduit à la méthode graphique suivante:

Représentons chaque mesure par un point de coordonnées rectangulaires (t_i, x_i) , divisons tous ces points en deux moitiés par une droite orthogonale à l'axe des x et construisons les centres de gravité des points de chaque moitié. Alors, la droite joignant les centres de gravité représente la relation compensée entre x et t .

La droite en question passe évidemment aussi par le centre de gravité des deux moitiés, c'est-à-dire de l'ensemble de tous les points, comme l'exige la méthode des moments (voir les équations (29) et (25)).

Il serait inutile d'insister sur la généralisation détaillée de la construction au cas de mesures d'inégale précision, car cette généralisation est très claire. On pourrait peut-être seulement rappeler qu'il faut modifier aussi les nombres m_I, m_{II} conformément aux poids adoptés pour les points.

7. Exemples numériques. Par la suite, je vais traiter en détail deux exemples d'application effective de notre méthode qui nous permettront de justifier les résultats théoriques, trouvés aux nos 3 et 4.

Interpolation linéaire.

Je commence par l'application de la méthode, exposée au no 6, aux mesures que j'ai faites pour déterminer la dépendance de la conductibilité thermique du ciment pulvérulent de la température.²⁾ J'ai choisi ces mesures avec l'intention de donner un exemple où l'erreur quadratique moyenne change avec la valeur de la variable indépendante: La précision de mes mesures diminue un peu avec la température croissante, comme on peut le constater simplement à la figure 1. J'y ai représenté les mesures effectuées par 12 points: 1, 2, . . . , 12, en portant sur un graphique les x_i (conductibilité thermique en unités techniques) en ordonnées et les températures (en 0° C) t_i en abscisses. Les valeurs correspondantes sont réunies dans la deuxième et la troisième colonnes du tableau 1.

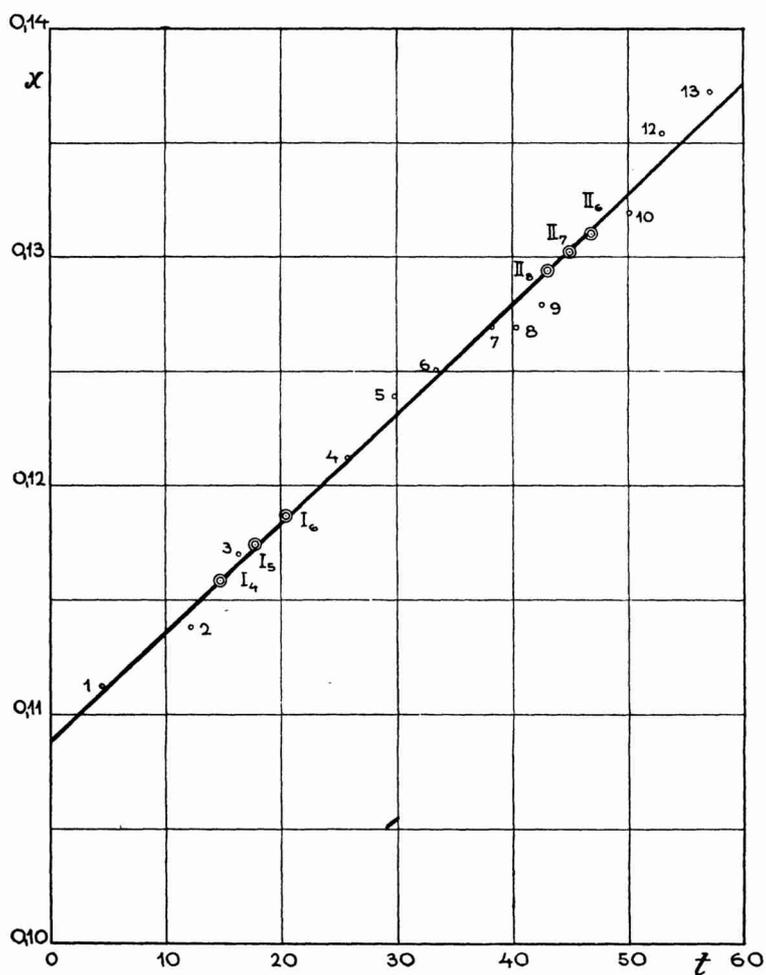
Je suppose que le phénomène étudié peut se représenter avec assez de précision par une fonction linéaire

$$x = A + Bt$$

²⁾ Technický Obzor, XLIV, p. 200—204, Praha 1936.

et je vais d'abord donner les résultats obtenus par la méthode des moments. Les valeurs les plus probables A_0, B_0 des paramètres A, B avec les erreurs quadratiques moyennes, résultant des calculs connus sur lesquels je n'insiste pas ici, sont les suivantes

$$A_0 = 0,1087 \pm 0,0006, B_0 = 0,000482 \pm 0,000017. \quad (35)$$



Appliquons d'abord rigoureusement notre méthode d'après la règle énoncée à la fin du numéro précédent, en choisissant les deux groupes comme suit:

1^{er} groupe: mesures: 1, 2, . . . , 6,
 2^e groupe: mesures: 7, 8, . . . , 12.

Alors les moyennes

$$t_I = 20,35, x_I = 0,11866,$$

$$t_{II} = 46,83, x_{II} = 0,13103$$

Tableau 1.

<i>i</i>	<i>t_i</i>	<i>x_i</i>	<i>m_I = m_{II} = 6</i>		<i>m_I = 5, m_{II} = 7</i>		<i>m_I = 4, m_{II} = 8</i>	
			<i>x'_i</i>	<i>ε_i</i>	<i>x'_i</i>	<i>ε_i</i>	<i>x'_i</i>	<i>ε_i</i>
1	4,5°	0,1112	0,1113	+0,0001	0,1112	0,0000	0,1109	-0,0003
2	12,1	0,1138	0,1148	+0,0010	0,1149	+0,0011	0,1145	+0,0007
3	16,4	0,1170	0,1168	-0,0002	0,1168	-0,0002	0,1166	-0,0004
4	25,8	0,1212	0,1212	0,0000	0,1212	0,0000	0,1211	-0,0001
5	29,9	0,1239	0,1231	-0,0008	0,1232	-0,0007	0,1231	-0,0008
6	33,4	0,1250	0,1248	-0,0002	0,1248	-0,0002	0,1248	-0,0002
7	38,3	0,1269	0,1270	+0,0001	0,1271	+0,0002	0,1271	+0,0002
8	40,4	0,1269	0,1279	+0,0010	0,1280	+0,0011	0,1281	+0,0012
9	42,6	0,1279	0,1290	+0,0011	0,1291	+0,0012	0,1292	+0,0013
10	50,1	0,1319	0,1325	+0,0006	0,1326	+0,0007	0,1328	+0,0009
11	52,9	0,1354	0,1339	-0,0015	0,1340	-0,0014	0,1342	-0,0012
12	56,7	0,1372	0,1354	-0,0016	0,1358	-0,0014	0,1360	-0,0012

et les équations (27) donnent les valeurs approchées des paramètres:

$$A_1 = 0,1092, B_1 = 0,000466. \quad (36)$$

En les portant dans les équations

$$x'_i = A_1 + B_1 t_i$$

on arrive aux valeurs résumées dans le tableau 1 où l'on trouve aussi les résidus correspondants ϵ_i . En tenant compte de (31) et (32), on obtient

$$\delta_I = 4,47 \cdot 10^{-4}, \delta_{II} = 12,2 \cdot 10^{-4}, \delta_{\bar{x}} = 7,6 \cdot 10^{-4}$$

et en vertu de (33), (34)

$$\delta_{A_1} = 7,5 \cdot 10^{-4}, \delta_{B_1} = 2,1 \cdot 10^{-5}. \quad (37)$$

On voit que les différences

$$A_0 - A_1 = -0,0005, B_0 - B_1 = +0,000016 \quad (38)$$

sont plus petites que les erreurs quadratiques moyennes, indiquées dans (35), et donc les valeurs (36) sont suffisamment précises.

Cependant, les nombres m_I, m_{II} ne satisfont pas à la relation théorique (16'), car le rapport

$$\frac{m_I}{m_{II}} : \frac{\delta_I}{\delta_{II}} = \frac{6}{6} : \frac{4,47}{12,2} = 2,73 \quad (39)$$

n'est pas égal à l'unité. On peut en conclure que la répartition choisie n'est pas la plus avantageuse. C'était évident a priori, puis que nous avons déjà constaté, sur la figure 1, la diminution de la précision des mesures avec la température croissante. Donc, en conséquence de cette constatation, on a du choisir, dès le début, m_I inférieur à m_{II} ce qui imposerait à prendre par exemple

$$m_I = 5, m_{II} = 7. \quad (40)$$

Nous aurons ainsi

$$\begin{aligned} t_I &= 17,74, x_I = 0,11742, \\ t_{II} &= 44,91, x_{II} = 0,13017 \end{aligned}$$

et des équations (27) découlent les valeurs

$$A_2 = 0,1091, B_2 = 0,000471 \quad (41)$$

plus voisines de A_0, B_0 que A_1, B_1 de sorte que les écarts

$$A_0 - A_2 = -0,0004, B_0 - B_2 = +0,000011$$

sont plus faibles que les écarts antérieurs (38). Les valeurs x'_i et les résidus ε_i , correspondant à (40), se trouvent dans le tableau 1. On en déduit

$$\delta_I = 6,60 \cdot 10^{-4}, \delta_{II} = 10,9 \cdot 10^{-4}$$

et en raison des relations (32) et (30)

$$\delta_{A_2} = 6,8 \cdot 10^{-4}, \delta_{B_2} = 1,9 \cdot 10^{-5}. \quad (42)$$

Il s'en suit

$$\frac{m_I}{m_{II}} : \frac{\delta_I}{\delta_{II}} = \frac{5}{7} : \frac{6,60}{10,9} = 1,18 \quad (43)$$

et l'on voit que ce rapport est plus voisin de 1 que la valeur (39) et que simultanément les erreurs des paramètres sont plus petites, ce à quoi il fallait s'attendre.

Cependant, même pour le deuxième choix des nombres m_I, m_{II} , le rapport $m_I : m_{II}$ est un peu trop fort. Il est donc plausible d'essayer d'obtenir, en le diminuant, un résultat encore plus précis. Nous poserons donc

$$m_I = 4, m_{II} = 8 \quad (44)$$

ce qui entraîne

$$\begin{aligned} t_I &= 14,70, x_I = 0,11578, \\ t_{II} &= 43,04, x_{II} = 0,12944, \end{aligned}$$

et donne les valeurs

$$A_3 = 0,1087, B_3 = 0,000482$$

qui coïncident, par hasard, avec les valeurs les plus probables données par la méthode des moments. En portant les valeurs x'_i et ε_i , empruntées au tableau 1, dans les expressions (30) à (32), nous aurons

$$\begin{aligned} \delta_I &= 5,0 \cdot 10^{-4}, \quad \delta_{II} = 10,4 \cdot 10^{-4}, \\ \delta_{A_3} &= 5,9 \cdot 10^{-4}, \quad \delta_{B_3} = 1,6 \cdot 10^{-5}, \end{aligned} \quad (45)$$

$$\frac{m_I}{m_{II}} : \frac{\delta_I}{\delta_{II}} = \frac{4}{8} : \frac{5,0}{10,4} = 1,04. \quad (46)$$

En effet, le choix (44) est le plus avantageux, parce qu'il conduit aux valeurs identiques à A_0, B_0 . En même temps, les erreurs (45) sont plus petites que les erreurs (37) et (42) et le rapport (46) est, au surplus, pratiquement égal à l'unité.

En somme, on peut dire que l'exemple, que je viens de traiter, correspond bien aux résultats théoriques déduits au no 3. Surtout il vérifie, dans le cas considéré, la condition générale (16) et montre que la méthode donne des résultats d'autant plus précis qu'on satisfait plus rigoureusement à cette condition. Il faut cependant souligner que toutes les répartitions adoptées offrent des résultats tout-à-fait satisfaisants puis que leurs écarts à partir des valeurs les plus probables sont inférieurs aux erreurs quadratiques moyennes.

Interpolation quadratique.

Comme deuxième exemple, je choisis le problème consistant à compenser les mesures des vitesses x_i de l'eau à diverses profondeurs t_i , traité par la méthode des moments par M. Čurík.³⁾ Comme nous le verrons, les mesures sont, dans ce cas-là, d'égale précision et je vais profiter de la simplicité du problème pour mettre en évidence, la manière dont la répartition qualitative des mesures par groupes influence l'exactitude des résultats obtenus par notre méthode.

Les valeurs données par les mesures sont réunies dans le tableau 2 où j'ai déjà rangé ces mesures selon la grandeur croissante des valeurs x_i . Si l'on suppose que la dépendance de la vitesse x de la profondeur t est de la forme

$$x = A + Bt + Ct^2,$$

l'application de la méthode des moments donne les valeurs des paramètres A, B, C les plus probables A_0, B_0, C_0 :

³⁾ Počet vyrovnávací, Praha 1936, p. 234—236.

$$\begin{aligned} A_0 &= 0,2807 \pm 0,0023, \\ B_0 &= 0,1511 \pm 0,0064, \\ C_0 &= -0,3125 \pm 0,0069 \end{aligned} \quad (47)$$

où l'on a ajouté les erreurs quadratiques moyennes, calculées d'après les formules connues en partant de l'erreur commune aux observations directes

$$\delta = \sqrt{\frac{[\varepsilon\varepsilon]}{m-p}} = 0,00170. \quad (48)$$

Nous appliquons notre méthode approximative, conformément à la règle énoncée au n° 6, en choisissant les groupes suivants:

$$\left. \begin{array}{l} 1^{\text{er}} \text{ groupe: mesures } 1, 2, \\ 2^{\text{e}} \text{ groupe: mesures } 3, 4, \\ 3^{\text{e}} \text{ groupe: mesures } 5, 6 \end{array} \right\} \quad (49)$$

Tableau 2.

i	t_i	x_i	t_i^2	$\bar{t}_{(r)}$	$\bar{t}_{(r)}^2$	$\bar{x}_{(r)}$	x'_i	ε_i
1	1,0	0,12	1,00	0,9	0,82	0,16	0,11875	-0,00125
2	0,8	0,20	0,64				0,20125	+0,00125
3	0,6	0,26	0,36	0,3	0,18	0,27	0,25875	-0,00125
4	0,0	0,28	0,00				0,28125	+0,00125
5	0,4	0,29	0,16	0,3	0,10	0,295	0,29125	+0,00125
6	0,2	0,30	0,04				0,29875	-0,00125

selon le numérotage indiqué dans le tableau 2. D'après le même tableau, les équations (22) pour les paramètres inconnus deviennent

$$\left. \begin{array}{l} 0,160 = A + 0,9B + 0,82C, \\ 0,270 = A + 0,3B + 0,18C, \\ 0,295 = A + 0,3B + 0,10C \end{array} \right\} \quad (50)$$

d'où l'on tire

$$A = 0,28125, \quad B = 0,1500, \quad C = -0,3125. \quad (51)$$

On voit que tous les résidus ε_i , indiqués au tableau 2, sont — en valeur absolue — exactement égaux. Il en est de même des erreurs $\delta_{(r)}$ et toutes les mesures sont donc d'égale précision. Il vient

$$\delta_{(1)} = \delta_{(2)} = \delta_{(3)} = \delta = \sqrt{\frac{[\varepsilon\varepsilon]}{m-p}} = 0,00176$$

ce qui montre, par rapport à (48), que les résultats obtenus par notre méthode sont à peu près aussi précis que ceux qu'on obtient par la méthode des moments. Les écarts

$$A_0 - A = -0,0005_5, B_0 - B = +0,0011, C_0 - C = 0,0000$$

sont, en effet, environ cinq fois plus petits que les erreurs à craindre, indiquées dans les équations (47). Le problème se trouve ainsi complètement résolu; je le reprendrai néanmoins pour étudier la dépendance de la précision des résultats obtenus du choix des groupes, sans changer les nombres $m_{(1)} = m_{(2)} = m_{(3)} = 2$. Outre le choix (49), fait en accord avec la règle générale, je vais considérer, dans ce but, encore cinq répartitions qualitativement différentes,

Tableau 3.

	Groupes			A	B	C
	1 ^{er}	2 ^e	3 ^e			
I	1 2	3 4	5 6	0,2807	0,1500	-0,3125
II	4 6	3 5	1 2	0,2807	0,1500	-0,3125
III	4 5	2 6	1 3	0,2791	0,1543	-0,3125
IV	1 4	3 6	2 5	0,2875	0,1375	-0,3125
V	1 6	2 5	3 4	0,2833	0,1310	-0,2924
VI	2 4	3 6	1 5	0,2777	0,1723	-0,3333
Méthode des moments				A_0 0,2807	B_0 0,1511	C_0 -0,3125

Tableau 3.

$A_0 - A$	$B_0 - B$	$C_0 - C$	$\bar{x}_{(2)} - \bar{x}_{(1)}$	$\bar{x}_{(3)} - \bar{x}_{(1)}$
-0,0005 ₅	+0,0011	0,0000	0,110	0,135
-0,0005 ₅	+0,0011	0,0000	0,115	0,130
+0,0016	-0,0032	0,0000	0,060	0,095
-0,0068	+0,0136	0,0000	0,045	0,080
-0,0026	+0,0201	-0,0201	0,035	0,060
+0,0030	-0,0212	+0,0208	0,035	0,075
erreurs quadratiques moyennes:				
0,0023	0,0064	0,0069		

indiquées dans le tableau 3. On y trouve les numéros des mesures contenues dans le premier, le second et le troisième groupe, les valeurs des paramètres obtenues par notre méthode appliquée à la