Sommes de Luhn généralisées

Dr Jean-Yves Degos

16 octobre 2025, révisé le 23 octobre 2025

Résumé

Cette note, à partir du cas d'une anomalie Flores, propose l'introduction de sommes de Luhn généralisées pour compléter la somme de Luhn usuelle associée à un SIRET ou un SIREN. Le but est de construire des invariants supplémentaires qui permettent de détecter les inversions de chiffres dans un SIRET, même si elles concernent des chiffres à des positions de même parité, ce qu'échoue à faire la somme de Luhn usuelle.

1 La somme de Luhn usuelle

Étant donné un SIRET dont les 14 chiffres sont notés :

$$a_{14}\,a_{13}\,a_{12}\,a_{11}\,a_{10}\,a_9\,a_8\,a_7\,a_6\,a_7\,a_6\,a_5\,a_4\,a_3\,a_2\,a_1\,,$$

la SOMME DE LUHN se calcule par la formule :

$$\sum_{i=1}^{7} a_{2i-1} + \sum_{i=1}^{7} r(2\times a_{2i})$$

où r désigne la réduction modulo 9. Si l'on a bien affaire à un SIRET, la somme de Luhn est un multiple de 10.

Plus précisément, lorsque a et b sont deux entiers compris entre 0 et 9, l'application r est définie ainsi :

$$r(10\times a+b) = \begin{cases} b & \text{si} \quad a=0\,, \\ b+a & \text{si} \quad a>0\,. \end{cases}$$

Exemple : le SIRET 215 704 636 00012 de la Commune de Metz a pour somme de Luhn :

$$2+0+0+3+4+7+1+r(2\times1)+r(2\times0)+r(2\times6)+r(2\times6)+r(2\times0)+r(2\times5)+r(2\times2)$$

soit :

$$17+r(2)+r(0)+r(12)+r(12)+r(0)+r(10)+r(4) = 17+2+0+3+3+0+1+4 = 17+13 = 40\,,$$

qui est bien un multiple de 10.

2 Une anomalie Flores

Durant la campagne 2023, nous avons été confrontés une anomalie impliquant les deux « SIRET » suivants :

- 446 034 290 00010 ;
- 443 064 290 00010.

On croit qu'il y a deux SIRET, mais en fait, seul le second code est celui d'un SIRET présent dans la base Sirene (ASSOCIATION DE COMMUNES ENFANCE JEUNESSE, 66 PLACE de la Mairie, 73100 Grésy-sur-Aix), car le premier code a en fait été identifié comme un numéro Adeli.

L'anomalie n'en était donc pas une, mais elle aura mis en difficulté trois êtres humains pendant environ trois heures pour cette raison fondamentale : la somme de Luhn est invariante par permutation de deux chiffres dont les positions ont même parité. C'est bien le cas ici, puisque les chiffres 6 et 3 qui correspondent aux positions 12 et 14 ont été permutés. Il en résulte que si la somme de Luhn est un multiple de 10 dans un cas, c'est également un multiple de 10 dans l'autre cas.

3 Les sommes de Luhn généralisées

On repart d'un SIRET dont les chiffres sont notés :

$$a_{14}\,a_{13}\,a_{12}\,a_{11}\,a_{10}\,a_9\,a_8\,a_7\,a_6\,a_7\,a_6\,a_5\,a_4\,a_3\,a_2\,a_1\,,$$

et on introduit un entier d compris entre 1 et 9.

On appelle d-SOMME DE LUHN la quantité :

$$s_d = \sum_{k=1}^{d} \sum_{i=1}^{s(k,d)} r(k \times a_{k+(i-1)d}) \text{ où } s(k,d) = 1 + \left\lfloor \frac{14-k}{d} \right\rfloor\,,$$

et où r désigne la réduction modulo 9. En effet, on doit toujours avoir :

$$1 \leq k + (i-1)d \leq 14\,.$$

La 2-somme de Luhn s_2 correspond évidemment à la somme de Luhn usuelle, et la 1-somme de Luhn est la somme usuelle. Mais lorsque que la somme de Luhn usuelle est identique pour deux « SIRET », on peut les distinguer à l'aide des sommes s_d pour $d > 2$.

3.1 Calcul de $s_3(446\ 034\ 290\ 00010)$

On a :

$$\begin{aligned}
s_3(446\ 034\ 290\ 00010) &= r(1\times 0)+r(1\times 0)+r(1\times 9)+r(1\times 3)+r(1\times 4)\\
&\quad +r(2\times 1)+r(2\times 0)+r(2\times 2)+r(2\times 0)+r(2\times 4)\\
&\quad +r(3\times 0)+r(3\times 0)+r(3\times 4)+r(3\times 6)\\
&= (0+0+9+3+4)+(2+0+4+0+8)+(0+0+3+9)\\
&= 16+14+12\\
&= 42\,.
\end{aligned}$$

3.2 Calcul de $s_3(443\ 064\ 290\ 00010)$

On a :

$$\begin{aligned}
s_3(443\ 036\ 290\ 00010) &= r(1\times 0)+r(1\times 0)+r(1\times 9)+r(1\times 6)+r(1\times 4)\\
&\quad +r(2\times 1)+r(2\times 0)+r(2\times 2)+r(2\times 0)+r(2\times 4)\\
&\quad +r(3\times 0)+r(3\times 0)+r(3\times 4)+r(3\times 3)\\
&= (0+0+9+6+4)+(2+0+4+0+8)+(0+0+3+9)\\
&= 19+14+12\\
&= 45\,.
\end{aligned}$$

Ainsi les deux « SIRET » 446 034 290 00010 et 443 064 290 00010 diffèrent-ils par leur 3-somme de Luhn.

4 Question ouverte

Oublions maintenant le contexte des SIRET, pour considérer des suites de chiffres compris entre 0 et 9, avec éventuellement $n > 14$:

$$a_n\ \ldots\ a_2\, a_1\,,$$

On définit :

$$s_d = \sum_{k=1}^{d} \sum_{i=1}^{s(k,d)} r(k \times a_{k+(i-1)d}) \text{ où } s(k,d) = 1 + \left\lfloor \frac{n-k}{d} \right\rfloor\,,$$

et on considère la suite :

$$s_n\ \ldots\ s_2\, s_1\,,$$

Question : Dans quelle mesure la suite $(s_d)_{1\leq d\leq n}$ caractérise-t-elle la suite $(a_d)_{1\leq d\leq n}$?