Sur Carl Ludwig Siegel

Associé étranger

par M. Jean Dieudonné (*)

Carl Ludwig Siegel est l'un des mathématiciens les plus éminents du xx^e siècle. Il est né à Berlin le 31 décembre 1896; son père était employé des Postes. Il commença ses études supérieures de 1915 à 1917 à l'Université de Berlin, où il suivit entre autres un cours de Frobenius qui lui fit connaître la Théorie des nombres. Mobilisé en 1917, il ne put s'adapter à la vie militaire et dut être réformé pour troubles mentaux. Il ne reprit ses études qu'en 1919-1920 à l'Université de Göttingen, où il passa sa thèse et son « Habilitation » sous la direction de E. Landau, un des spécialistes les plus estimés de théorie analytique des nombres. Il a été professeur à l'Université de Francfort de 1922 à 1937, et à l'Université de Göttingen de 1938 à 1940. Invité en 1940 à faire des conférences au Danemark et en Norvège, il en profita pour quitter la Norvège quelques jours avant l'invasion allemande et chercher refuge aux États-Unis pour échapper au régime nazi qu'il exécrait. Il avait déjà été invité à Princeton pour une année en 1935; il y travailla à l'Institute for Advanced Study de 1940 à 1951 et y fut nommé professeur permanent en 1946. En 1951, il reprit sa chaire à Göttingen, où il a résidé jusqu'à sa mort le 4 avril 1981.

La renommée universelle de Siegel est surtout due à ses travaux de Théorie des nombres, où il s'inscrit dans la grande lignée qui commence avec Fermat et se poursuit avec Euler, Lagrange, Gauss et les brillantes écoles allemande et française du xix^e siècle. Mais on lui doit aussi d'importants résultats en théorie des fonctions de plusieurs variables complexes et en mécanique céleste; il est d'ailleurs frappant que tous ses mémoires de Théorie des nombres reposent sur un maniement de l'Analyse mathématique d'une profondeur et d'une virtuosité incomparables. Ses démonstrations se caractérisent par une puissance impressionnante : une fois en possession d'une idée, Siegel la développe jusqu'au bout, quel que soit le maquis de formules ou de difficultés techniques qu'il rencontre sur son chemin. De ce fait, un mémoire de Siegel est difficilement améliorable par la méthode même de Siegel (ce qui explique qu'il ait eu peu d'élèves directs); pour aller plus loin, il faut procéder autrement. Il se dégage aussi de son œuvre un souci constant de perfection : tous les énoncés arithmétiques sont poussés jusqu'à leurs conclusions numériques, avec de nombreux exemples qui les vérifient. Le nombre d'erreurs (soit de calcul, soit de démonstration) est remarquablement faible; utiliser des formules de Siegel est un plaisir : on sait qu'elles sont vraies. Il s'attachait à cette perfection (un peu décourageante pour ceux qui auraient voulu le suivre) jusque dans ses cours, qu'il faisait sans une note ni une hésitation, écrivant au tableau de son écriture parfaitement régulière des formules d'une effroyable complexité sans la moindre erreur.

(*) Notre confrère J.-P. Serre m'a apporté une aide précieuse dans la description et l'évaluation des travaux de Siegel.

Ayant voué toute sa vie à la recherche mathématique (il ne s'était jamais marié), il n'avait cependant rien d'un ascète et aimait voyager; il a, par exemple, donné de nombreux cours au Tata Institute de Bombay. Il souffrait difficilement la médiocrité, et ses jugements sur les autres mathématiciens, qu'il exprimait parfois avec un humour féroce, n'étaient souvent pas très charitables. Il n'a cessé de travailler tant que sa santé le lui a permis, et il a joui d'une longévité scientifique exceptionnelle, ayant encore publié entre 70 et 80 ans toute une série d'importants mémoires qui ont été réunis en un volume supplémentaire de ses œuvres.

La thèse de Siegel (1920) sort déjà tout à fait de l'ordinaire, et a été un jalon important dans la théorie des approximations diophantiennes. Liouville avait le premier remarqué que les nombres algébriques de degré >1 ont une « mauvaise » approximation par des nombres rationnels : si ξ est algébrique de degré $n>1$, il y a une constante $C(\xi)$ telle que, pour tout nombre rationnel irréductible p/q, on ait :

(1) $$\left|\xi-\frac{p}{q}\right|>C(\xi)/q^n.$$

La démonstration de ce résultat est immédiate. C'est seulement en 1908, par une méthode beaucoup plus élaborée, que A. Thue put améliorer ce théorème en prouvant que l'inégalité :

(2) $$\left|\xi-\frac{p}{q}\right|\leqq\frac{1}{q^{n/2+1+\varepsilon}},$$

(pour un $\varepsilon>0$ quelconque) n'est possible que pour un nombre *fini* de valeurs de p/q. C'est ce résultat que Siegel améliora encore de façon significative en montrant que la même conclusion est valable pour l'inégalité :

(3) $$\left|\xi-\frac{p}{q}\right|\leqq\frac{1}{q^{2\sqrt{n}}}.$$

Sa démonstration, extrêmement ingénieuse, procède comme celle de Thue par l'absurde. En fait, il considère au lieu de (3) l'inégalité analogue où $2\sqrt{n}$ est remplacé par :

(4) $$\beta=\frac{n}{s+1}+s+\frac{1}{2},$$

où s est un entier *quelconque* tel que $0\leqq s\leqq n-1$. L'hypothèse que l'inégalité :

(5) $$\left|\xi-\frac{p}{q}\right|\leqq\frac{1}{q^{\beta}},$$

a une infinité de solutions permet d'en considérer deux, p_1/q_1 et p_2/q_2, où q_1 et $r=[\log q_2/\log q_1]$ sont *arbitrairement grands*. Le point essentiel de la preuve consiste à montrer l'existence d'une constante C ne dépendant que de ξ telle qu'il existe un entier $\rho<(r/16n)+n^2$ pour lequel un des deux nombres :

$$E_1=C^r q_1^{m+r}q_2^s\left|\xi-\frac{p_1}{q_1}\right|^{r-\rho},\qquad E_2=C^r q_1^{m+r}q_2^s\left|\xi-\frac{p_2}{q_2}\right|,$$

est $\geqq 1$ (avec $m=[((n+(1/16n))/(s+1)-1)r]$). Pour cela, Siegel construit (ce qui est la difficulté de la preuve) une famille de polynômes $R_\rho(x, y)$ (pour toute valeur de l'entier $\rho<r$) à coefficients entiers $\leqq C'^r$ (où C′ ne dépend que de ξ), de degrés $m+r-\rho$ en x, s en y, et prouve (par la méthode des tiroirs de Dirichlet) qu'il y a une valeur de ρ pour laquelle :

1° le nombre :

$$q_1^{m+r-\rho} q_2^s R_\rho\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)$$

est $\neq 0$, donc de valeur absolue $\geqq 1$ puisque c'est un entier;

2° on a :

$$\left|R_\rho\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)\right| \leqq C''^r \left(1+\left|\frac{p_1}{q_1}\right|\right)^{m+r} \left(1+\left|\frac{p_2}{q_2}\right|\right)^s \left(\left|\xi-\frac{p_1}{q_1}\right|^{r-\rho}+\left|\xi-\frac{p_2}{q_2}\right|\right),$$

où C″ ne dépend que de ξ. L'une des relations $E_1 \geqq 1$, $E_2 \geqq 1$ en découle.

D'autre part, on montre aisément que :

$$(m+r)\log q_1 + r \log C < (\beta - s) \log q_2,$$

et l'hypothèse que p_1/q_1 et p_2/q_2 vérifient tous deux l'inégalité (5) entraîne alors que $E_1 < 1$ et $E_2 < 1$, ce qui est la contradiction cherchée.

A part une courte note de 1926, Siegel ne revient plus entre 1922 et 1929 aux questions d'approximation diophantienne. Son « Habilitationsschrift » de 1922 marque son premier contact avec les formes quadratiques, dans l'esprit des travaux de Hardy-Littlewood et de Landau sur les sommes de carrés. Le silence qu'il observe ensuite est certainement occupé par l'élaboration du long mémoire qu'il publie en 1929 dans les *Abhandlungen* de Berlin, sous le titre neutre de « Quelques applications des approximations diophantiennes », sans doute son œuvre la plus profonde et la plus difficile, et qui va aussitôt le rendre célèbre.

Ce mémoire comprend deux parties très différentes et qui ne sont reliées que par un fil très ténu. La seconde est celle dont le retentissement fut le plus grand, car elle résout un problème fondamental de la théorie des équations diophantiennes : donner des conditions nécessaires et suffisantes pour qu'une équation polynomiale à coefficients entiers :

$$f(x, y)=0, \tag{6}$$

n'ait qu'un nombre *fini* de solutions (x, y) en nombres *entiers*. Jusque-là le meilleur résultat était celui de Thue, suivant lequel une équation (6) de la forme $g(x, y)-a=0$, où $a\neq 0$ et g est homogène et de degré $\geqq 3$, n'a qu'un nombre fini de solutions entières. Le théorème de Siegel a un énoncé très simple : le *seul* cas où, pour un polynôme irréductible f, l'équation (6) ait une infinité de solutions entières est celui où la courbe Γ d'équation (6) admet une représentation paramétrique par des fonctions rationnelles dont le dénominateur est un polynôme du premier ou du second degré (ce qui entraîne que Γ est de genre 0 et a au plus deux points à l'infini). Nous nous bornerons à esquisser la démonstration lorsque Γ est de genre 1.

Siegel commence par prouver un résultat général valable pour toute équation (6). Soit L le corps des fonctions rationnelles sur la courbe Γ, et soit g une fonction de ce corps d'ordre m (nombre qui par la suite sera pris arbitrairement grand). On suppose que pour une infinité de valeurs de x/z et y/z, où x, y, z sont entiers premiers entre eux et $f(x/z, y/z)=0$, $g(x/z, y/z)$ a une valeur entière. Alors il y a une suite extraite de l'ensemble des points $(x/z, y/z)$ précédents, qui converge vers un point de Γ qui est un *pôle* de g ; et si r est l'ordre de ce pôle, pour toute fonction φ du corps L qui s'annule en ce pôle et tout $\varepsilon>0$, il y a une constante $C(\varphi, \varepsilon)$ telle que :

$$|\varphi(x/z, y/z)| \leqq C(\varphi, \varepsilon)\,(|x|+|y|+|z|)^{-(m/hr)+\varepsilon}, \tag{7}$$

pour tout point $(x/z, y/z)$ de la suite considérée (h étant le degré de f).

On suppose alors que Γ est de genre 1, donc admet une représentation paramétrique par des fonctions elliptiques, soit $x=w(s)$, $y=v(s)$; soit r le nombre de racines d'une équation $w(s)=a$ dans un parallélogramme de périodes. L'idée essentielle est d'utiliser le théorème de Mordell affirmant que, si M est le Z-module des nombres complexes tels que $w(s)$ et $v(s)$ soient tous deux *rationnels*, M admet un système *fini* de générateurs $s_1, \ldots, s_q$. Soit n un entier fixé (qui sera pris par la suite arbitrairement grand) ; par division euclidienne, tout élément de M peut s'écrire :

$$s=n\sigma+c, \tag{8}$$

où $\sigma\in M$, et $c\in M$ ne prend qu'un nombre *fini* de valeurs. Le théorème se démontre alors (comme celui de Thue) par l'absurde, en supposant donc que l'équation (6) a une infinité de solutions entières ; il y a donc aussi une infinité de ces solutions pour lesquelles le nombre $c\in M$ a la même valeur. On va appliquer le résultat (7) à la fonction rationnelle $g(x/z, y/z)=x$; par le théorème d'addition des fonctions elliptiques, $w(ns+c)$ est une fonction appartenant au corps L, d'ordre n^2m, dont les n^2m pôles sont au plus d'ordre m ; ces pôles ont des coordonnées qui sont des nombres algébriques de degré $\leqq n^2m$. D'après (7), un de ces pôles est limite d'une suite de points $(\xi/\zeta, \eta/\zeta)$ de Γ, où ξ, η, ζ sont des entiers premiers entre eux. Si la limite ρ de la suite des ξ/ζ est finie, c'est un nombre algébrique de degré $\leqq n^2m$, et on a d'après (7) :

$$\left|\frac{\xi}{\zeta}-\rho\right| \leqq C(|\xi|+|\zeta|)^{-kn^2m}, \tag{9}$$

où $k>0$ est indépendant de n. D'autre part, l'inégalité (3) de Siegel montre que l'on a, sauf pour un nombre fini de nombres ξ/ζ de la suite précédente :

$$\left|\frac{\xi}{\zeta}-\rho\right| \geqq C'(|\xi|+|\zeta|)^{-n\sqrt{m}}, \tag{10}$$

où la constante C' ne dépend que de n. Mais il est clair que pour n assez grand les deux inégalités (9) et (10) ne peuvent être simultanément vérifiées que par un nombre fini de couples d'entiers (ξ, ζ), ce qui amène la contradiction cherchée. Le raisonnement est encore plus simple (en considérant la suite des ζ/ξ) lorsque $|\xi/\zeta|$ tend vers $+\infty$.

Lorsque Γ est de genre $\geqq 2$, Siegel utilise comme outil principal la généralisation, due à A. Weil, du théorème de Mordell, où il faut cette fois se placer dans la jacobienne de Γ (ou, comme le fait Siegel, considérer sa paramétrisation par les fonctions Θ de Riemann). Mais pour pouvoir l'appliquer comme ci-dessus, il faut une suite de constructions d'une extrême complication.

Bien que demeurant difficiles, les démonstrations du théorème de Siegel que l'on connaît aujourd'hui se sont un peu simplifiées. Son résultat sur les approximations diophantiennes des nombres algébriques a été amélioré par Roth, qui a montré que l'on peut y remplacer (3) par :

$$\left|\xi-\frac{p}{q}\right| \leqq \frac{1}{q^{2+\varepsilon}}, \tag{11}$$

pour tout $\varepsilon>0$. L'exposant 2 est cette fois le meilleur possible; toutefois on est encore hors d'état de donner une expression effectivement calculable en fonction de ξ d'une majoration pour les entiers q vérifiant même l'inégalité moins bonne (2) de Thue. De même, sauf lorsque la courbe (6) est de genre 1, ou pour certaines équations particulières de genre plus grand, on ne sait pas majorer effectivement les solutions de (6) en fonction des coefficients de l'équation.

La première partie du mémoire de Siegel de 1929 ne le cède en rien à la seconde pour l'originalité et la puissance imaginative. Avec la preuve de la transcendance de e^{π} par Gelfond, publiée la même année, c'est le premier résultat nouveau sur les nombres transcendants depuis les fameux théorèmes de Hermite et Lindemann sur la transcendance de e et π, qui remontaient à près d'un demi-siècle : Siegel prouve en effet que, pour la fonction de Bessel J_0, $J_0(\xi)$ est transcendant pour tout nombre algébrique réel $\xi \neq 0$. Cela résulte de la proposition beaucoup plus précise suivante : si $g(y, z)$ est un polynôme de degré total $p>0$ dont les coefficients sont des entiers de valeur absolue $\leqq G$, et ξ un nombre algébrique réel $\neq 0$ et de degré m, on a :

$$|g(J_0(\xi), J_0'(\xi))| \geqq c\, G^{-123\, p^2 m^2}, \tag{12}$$

où c ne dépend que de p et ξ.

La méthode de Siegel est tout à fait nouvelle. Il commence par une étude analytique, dans l'esprit de Liouville, sur les relations algébriques entre x, $y=J_0(x)$ et $y'=J_0'(x)$, aboutissant au résultat suivant : soit :

$$\varphi(x)=\sum_{\beta} \sum_{\alpha=0}^{\beta} f_{\alpha\beta}(x)\, y(x)^{\alpha} y'(x)^{\beta-\alpha}, \tag{13}$$

où les $f_{\alpha\beta}$ sont des polynômes en x à coefficients réels, et non nuls pour q des monômes $y^{\alpha} y'^{\beta-\alpha}$; alors φ et ses dérivées jusqu'à l'ordre $q-1$ sont q formes linéaires en ces monômes, dont le déterminant est un polynôme en x non nul.

Soit l un entier $>p$, et en posant $k=(l+1)(l+2)/2$, soient $n \geqq 2k^2$ un entier arbitraire et $\varepsilon>0$ un nombre réel arbitraire. L'étape suivante est de former une fonction (13) telle que :

1° les k polynômes $f_{\alpha\beta}$ pour $\beta \leqq l$ sont de degré $\leqq 2n-1$, avec des coefficients entiers majorés par $(n!)^{2+\varepsilon}$;

2° le développement en série entière de φ commence par un terme en $x^{(2k-1)n}$, et admet une majorante, produit d'une constante indépendante de n, et de la série :

$$(n!) \sum_{\nu=(2k-1)n}^{\infty} \frac{|x|^{\nu}}{(\nu!)^{1-\varepsilon}}. \tag{14}$$

Notons $T_j(x)$ $(1 \leqq j \leqq k)$ les fonctions $J_0(x)^\alpha J_0'(x)^{\beta-\alpha}$ pour $\beta \leqq l$ et $\alpha \leqq \beta$, avec $T_1 = 1$. En raison de l'équation différentielle de J_0, on peut écrire pour la fonction φ précédente et tout ordre de dérivation $a \geqq 0$,

$$x^a \varphi^{(a)}(x) = \sigma_{a1} T_1(x) + \ldots + \sigma_{ak} T_k(x), \tag{15}$$

où les $\sigma_{ab}(x)$ sont des polynômes de degré $2n+a-1$, à coefficients entiers et $O((n!)^{3+2\varepsilon})$ pour $a < n+k^2$. Le point crucial consiste à montrer que pour $a \leqq n+k^2-1$ et tout nombre réel $\xi \neq 0$, la matrice $(\sigma_{aj}(\xi))$ est *de rang k*.

On peut alors prendre k entiers $h_\nu \leqq n+k^2-1$ tels que les k fonctions :

$$\varphi_\nu(x) = \sigma_{h_\nu 1}(\xi) T_1(x) + \ldots + \sigma_{h_\nu k}(\xi) T_k(x), \tag{16}$$

soient linéairement indépendantes. Soient $r = l-p$, $v = (r+1)(r+2)/2$, et considérons les v fonctions :

$$J_0(x)^\rho J_0'(x)^\sigma g(J_0(x)) \qquad \text{pour} \quad \rho+\sigma \leqq r, \tag{17}$$

qui s'écrivent :

$$\psi_\mu(x) = c_{\mu 1} T_1(x) + \ldots + c_{\mu k} T_k(x), \qquad (1 \leqq \mu \leqq v), \tag{18}$$

où les $c_{\mu j}$ sont entiers et $\leqq G$ en valeur absolue, et les v fonctions ψ_μ sont linéairement indépendantes; on le complète par $w = k - v$ des fonctions φ_ν, de façon à avoir k combinaisons linéairement indépendantes des T_j. On montre alors que le déterminant Δ des coefficients de ces k combinaisons est un polynôme en ξ de degré $w(3n+k^2-2)$, à coefficients entiers, qui sont $O((n!)^{3w+2\varepsilon Gv})$. Finalement on peut en conclure que $|\Delta|$ est majoré par :

$$K(n!)^{3w+2\varepsilon Gv}\left(\frac{g(J_0(\xi), J_0'(\xi))}{G} + (n!)^{1-2k}\right), \tag{19}$$

où $K \geqq 1$ est indépendant de n. Si maintenant $\xi \neq 0$ est un nombre *algébrique* de degré m, et c un entier rationnel tel que $c\xi$ soit entier algébrique, $c^{w(3n+k^2-2)}\Delta$ est un *entier algébrique* $\neq 0$. En écrivant que sa norme est $\geqq 1$, et en choisissant convenablement n en fonction de G, on obtient enfin (12).

Au cours de cette démonstration, Siegel avait introduit une famille de séries entières, qu'il appela E-fonctions; elles sont définies par des conditions arithmétiques sur leurs coefficients, et généralisent la fonction J_0. Plus tard Shidlovski parvint à généraliser à ces fonctions le théorème de Siegel (1954).

Siegel publia encore en 1932 un résultat sur les nombres transcendants : en adaptant la méthode de Gelfond (basée sur la formule d'interpolation de Newton) pour la preuve de la transcendance de e^π, il put montrer que, si les invariants g_2 et g_3 de la fonction elliptique $\wp(x)$ de Weierstrass sont des nombres algébriques, alors une au moins des périodes de $\wp(x)$ est un nombre transcendant.

De 1932 date aussi une étonnante reconstitution par Siegel d'une formule de Riemann, d'après une page de brouillon conservée à la bibliothèque de Göttingen et difficilement déchiffrable; elle permet un calcul facile de la fonction zêta sur la droite critique $\mathrm{Re}(s) = 1/2$. Elle a notamment été utilisée par N. Levinson en 1974 pour montrer que plus de 34% des zéros de $\zeta(s)$ sont sur cette droite et sont des zéros simples.

En 1934, H. Heilbronn était parvenu à prouver une conjecture de Gauss : le nombre de classes $h(-d)$ d'idéaux d'un corps quadratique imaginaire de discriminant $-d$ tend vers $+\infty$ avec d. Siegel, en 1935, par une utilisation habile de la relation entre les fonctions zêta de deux corps quadratiques et la fonction zêta de leur composé (extension de Q de degré 4) améliora considérablement ce résultat en montrant que lorsque d tend vers $+\infty$, on a :

$$\log h(-d) \sim \frac{1}{2} \log d.$$

Ce résultat fut plus tard généralisé par R. Brauer à des corps de nombres galoisiens de degré quelconque.

Là encore, on ne dispose pas d'une expression effective en fonction de d qui minorerait $h(-d)$, et par exemple on ne connaît pas tous les corps quadratiques pour lesquels $h(-d)=3$.

A partir de 1935, la plupart des travaux de Siegel en théorie des nombres vont être centrés sur la théorie arithmétique des formes quadratiques à coefficients entiers, à un nombre quelconque n de variables. Inaugurée par Lagrange et Gauss pour $n=2$ et $n=3$, cette théorie s'était développée au XX^{e} siècle avec les travaux de Legendre, Eisenstein, Hermite, H. S. Smith et Minkowski. L'œuvre de Siegel en représente en un certain sens l'achèvement; mais en même temps il va l'élargir considérablement et préparer ses extensions modernes en la mettant en rapport avec la théorie des groupes de Lie et des fonctions automorphes de plusieurs variables.

Cette période s'ouvre avec trois longs mémoires (193 pages en tout!) publiés de 1935 à 1937 aux *Annals of Mathematics*. Dans cette œuvre monumentale par son ampleur et sa complexité, Siegel traite du problème général de la transformation d'une forme quadratique Q à m variables et à coefficients entiers, en une forme analogue R à $n \leq m$ variables, au moyen de substitutions linéaires à coefficients entiers. Ceci s'exprime commodément en lagage matriciel : si S, T sont deux matrices symétriques d'ordres respectifs m, n, à coefficients entiers, il s'agit d'étudier les matrices X de type $m \times n$, à coefficients entiers, vérifiant l'équation :

$$(20) \qquad {}^tX.S.X=T.$$

Le premier mémoire est consacré au cas où S et T sont définies positives, pour lequel la théorie avait fait le plus de progrès. Le nombre $A(S, T)$ des matrices X vérifiant (20) est alors *fini*; lorsque $S=T$, on pose $A(S, S)=E(S)$, ordre du sous-groupe de $GL(m, \mathbb{Z})$ laissant S invariante (« unités » de S). Depuis Gauss, on introduit les notions de *classe* et de *genre* pour les formes quadratiques à coefficients entiers (ou les matrices symétriques correspondantes) : deux matrices symétriques S, S_1 à coefficients entiers sont dans la même classe s'il existe une matrice inversible Y à coefficients entiers telle que :

$$(21) \qquad {}^tY.S.Y=S_1;$$

lorsque $A(S, T)$ est fini, il ne dépend que des classes de S et T. Les définitions du genre données par Gauss (pour les formes binaires) et généralisées par Eisenstein et Smith, à l'aide de « systèmes de caractères », sont compliquées; Poincaré et Minkowski en ont indépendamment trouvé une définition plus simple : S et S_1 sont dans le même genre si d'une part l'équation (21) admet une solution Y à coefficients réels, et si d'autre part, pour tout entier q, il existe une matrice Y_q à coefficients entiers et de déterminant inversible mod. q, telle que :

$$ {}^tY_q . S . Y_q \equiv S_1 \qquad (\text{mod}.\, q). \tag{22} $$

Les méthodes de réduction d'Hermite montrent que, pour les formes définies positives, un genre ne contient qu'un nombre fini de classes. Considérons dans ce cas un genre donné contenant h classes, dans chacune desquelles on choisit un représentant $S_j (1 \leqq j \leqq h)$. Eisenstein et Smith associent au genre considéré sa « masse » :

$$ \frac{1}{\mathrm{E}(S_1)} + \ldots + \frac{1}{\mathrm{E}(S_h)}, \tag{23} $$

et Smith était parvenu à en trouver une expression explicite à l'aide des « caractères » du genre (retrouvée indépendamment par Minkowski). Le premier mémoire de Siegel étudie plus généralement, pour deux matrices S, T définies positives, le quotient :

$$ \mathrm{M}(S, T) = \frac{(\mathrm{A}(S_1, T)/\mathrm{E}(S_1)) + (\mathrm{A}(S_2, T)/\mathrm{E}(S_2)) + \ldots + (\mathrm{A}(S_h, T)/\mathrm{E}(S_h))}{(1/\mathrm{E}(S_1)) + \ldots + (1/\mathrm{E}(S_h))}, \tag{24} $$

où $S_1, \ldots, S_h$ sont les représentants du genre auquel appartient S ; il en donne l'expression sous forme de produit infini :

$$ \mathrm{M}(S, T) = \mathrm{A}_\infty(S, T) \prod_p d_p(S, T). \tag{25} $$

Dans cette formule p parcourt l'ensemble des nombres premiers; on note $\mathrm{A}_q(S, T)$ le nombre de solutions de la congruence :

$$ {}^tX . S . X \equiv T \,(\text{mod}.\, q) \tag{26} $$

en matrices entières modulo q; $d_p(S, T)$ est la limite de $\mathrm{A}_q(S, T)/q^{mn-(n(n+1)/2)}$ lorsque q parcourt la suite des puissances p^{N} de p (« valeur moyenne p-adique » de $\mathrm{A}(S, T)$). Enfin $\mathrm{A}_\infty(S, T)$ est une « valeur moyenne » analogue : pour $m > n$ et $m \neq 2$, c'est la limite du rapport du volume dans $\mathbb{R}^{mn}$ de l'image réciproque par $X \mapsto {}^tX . S . X$ d'un voisinage V de T dans l'espace des matrices symétriques positives (ouvert de $\mathbb{R}^{n(n+1)/2}$), au volume de V dans cet espace, lorsque V tend vers le point T. La démonstration procède par récurrence sur m et n, et est une adaptation très subtile des méthodes de Gauss, Dirichlet et Minkowski.

Les deux autres mémoires de cette série sont encore plus originaux ; le second traite des formes quadratiques « indéfinies » de signature quelconque. Siegel montre d'abord que le second membre de la formule (25) garde un sens sauf dans deux cas particuliers (celui où $m=2$ et $-\det(S)$ est un carré, et celui où $m-n=2$ et $-\det(ST)$ est un carré). Mais les formules (23) et (24) n'ont plus de sens parce que le sous-groupe de $\mathrm{GL}(m, \mathbb{Z})$ laissant invariante S est alors infini. Trouver ce qui devait remplacer le premier membre de (25) dans ce cas était un problème qui n'avait été abordé que dans un cas particulier par G. Humbert, pour les formes ternaires. Siegel le résout de façon générale : il considère, dans l'espace des matrices symétriques d'ordre m et de signature donnée, un voisinage ouvert B d'une matrice S, et son image réciproque B_1 dans l'espace $\mathbb{R}^{mn}$ par l'application $X \mapsto {}^tX . S . X$; B_1 est invariante par le groupe des « unités » de S agissant par multiplication à gauche; il y a un domaine fondamental D pour cette action dans B_1, et si le volume $v(B)$ est fini, il en est de même du volume $v(D)$, et la limite :

$$\rho(S) = \lim v(D)/v(B)$$

existe lorsque B tend vers S. Un genre donné auquel appartient S ne contient encore qu'un nombre fini de classes, et si $S_1, \ldots, S_h$ sont des représentants de ces classes, le nombre :

$$\mu(S) = \rho(S_1) + \rho(S_2) + \ldots + \rho(S_h) \tag{27}$$

est ce qui remplace la « masse » du genre de S. Il y a une définition analogue (plus compliquée) d'un nombre $\mu(S, T)$ qui remplace le numérateur de (24), et finalement la formule de Siegel (25) est valable en remplaçant le premier membre car $\mu(S, T)/\mu(S)$. Siegel devait encore améliorer ce résultat en 1944 en montrant que dans certains cas les termes figurant dans l'expression de $\mu(S, T)$ sont les mêmes pour toutes les classes d'un même genre.

Dans le troisième mémoire (1937), Siegel aborde la théorie des formes quadratiques à coefficients dans un corps de nombres algébriques; là, il n'avait aucun prédécesseur, et comme on peut s'y attendre, de nouvelles complications apparaissent, dont il triomphe avec sa maîtrise coutumière.

Revenant aux formes définies positives, lorsqu'un genre ne comporte qu'*une seule classe*, le premier membre de (25) se réduit à $A(S, T)$; lorsque S est la matrice unité, cela se produit pour $m \leqq 8$; si en outre $n = 1$, on retrouve ainsi toutes les formules données par Jacobi, Eisenstein, Smith et Liouville pour le nombre de représentations d'un entier comme somme de m carrés pour $4 \leqq m \leqq 8$. La démonstration de Jacobi pour $m = 4$ reposait sur son étude des fonctions thêta et de leurs relations avec le groupe modulaire $SL(2, \mathbf{Z})$, provenant de la formule (trouvée indépendamment par Gauss, Cauchy et Poisson) :

$$\theta(z) = \frac{1}{\sqrt{z}}\theta\left(\frac{1}{z}\right) \tag{28}$$

pour la plus simple des fonctions thêta :

$$\theta(z) = \sum_{n=-\infty}^{+\infty} e^{-\pi n^2 z} \qquad (\text{où Re } z > 0). \tag{29}$$

Dans son premier mémoire sur les formes quadratiques définies positives, Siegel remarque que sa formule (25) équivaut à une remarquable identité entre fonctions qui généralisent les formes modulaires. Il prend comme espace des variables ce que l'on appelle maintenant le « demi-espace de Siegel », généralisation du « demi-plan de Poincaré » classique : c'est l'ensemble des matrices symétriques *complexes* Z d'ordre n, dont les parties imaginaires sont des matrices symétriques définies positives. On considère (pour une matrice symétrique S définie positive à coefficients entiers) la fonction de Z généralisant les fonctions thêta :

$$f(S, Z) = \sum_{C} \exp(\pi i . \operatorname{Tr}({}^t C . S . C . Z)), \tag{30}$$

où C parcourt l'ensemble $\mathbb{Z}^{mn}$ des matrices entières de type $m \times n$; cette série converge absolument lorsque Z varie dans le demi-espace de Siegel. Il est facile de voir que l'on a :

$$f(S, Z) = \sum_{T} \mathrm{A}(S, T) \exp(\pi i . \mathrm{Tr}(TZ)), \tag{31}$$

où T parcourt l'ensemble des matrices symétriques entières d'ordre n. Cela étant, si l'on pose :

$$\mathrm{F}(S, Z) = \frac{(f(S_1, Z)/\mathrm{E}(S_1)) + \ldots + (f(S_h, Z)/\mathrm{E}(S_h))}{(1/\mathrm{E}(S_1)) + \ldots + (1/\mathrm{E}(S_h))}, \tag{32}$$

la formule (25) équivaut (pour m assez grand) à une expression de F (S, Z) sous la forme d'une série convergente :

$$\mathrm{F}(S, Z) = \sum_{K, L} c_{K, L} . \det(KZ + L)^{-m/2}, \tag{33}$$

où le couple (K, L) parcourt un certain ensemble de matrices entières d'ordre n, défini par des conditions arithmétiques. L'analogie évidente de la série (33) avec les séries d'Eisenstein (qui correspondent au cas $n = 1$) incita Siegel, dans plusieurs mémoires ultérieurs, à faire une étude systématique de ce qu'il appela les *formes modulaires de degré n.* Il s'agit de fonctions holomorphes définies dans le demi-espace de Siegel D; sur cet espace, le *groupe symplectique* $\mathrm{Sp}(2n, \mathbb{R})$, formé des matrices d'ordre $2n$:

$$U = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

telles que ${}^tU . J . U = J$ pour la matrice $\mathrm{J} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$, agit par :

$$Z \mapsto (AZ + B)(CZ + D)^{-1}.$$

Dans un premier mémoire de 1939, Siegel considère le sous-groupe $\mathrm{Sp}(2n, \mathbb{Z})$ de $\mathrm{Sp}(2n, \mathbb{R})$, qui n'est autre que le groupe des transformations des systèmes de $2n$ périodes d'un système de n intégrales abéliennes de première espèce linéairement indépendantes sur une surface de Riemann de genre n. Ce groupe agit de façon proprement discontinue sur D, et Siegel en détermina un domaine fondamental à l'aide de la réduction de Minkowski des formes quadratiques. Les formes modulaires de degré n et de poids r sont les fonctions holomorphes dans D, qui se transforment par le groupe $\mathrm{Sp}(2n, \mathbb{Z})$ suivant la loi :

$$\psi_r((AZ + B)(CZ + D)^{-1}) = \det(CZ + D)^r \psi_r(Z) \quad (r \text{ pair}). \tag{34}$$

Siegel obtient pour ces formes une expression par des séries généralisant les séries d'Eisenstein; de là il passa naturellement aux *fonctions modulaires*, fonctions méromorphes dans D invariantes par le groupe $\mathrm{Sp}(2n, \mathbb{Z})$; on en obtient par exemple en prenant le quotient de deux formes modulaires de même poids, et plus tard (1960) Siegel prouva qu'on les obtient toutes de cette façon; il montra de plus que l'ensemble des fonctions modulaires est un corps dont le degré de transcendance sur $\mathbb{C}$ est $n(n+1)/2$.

Il inaugurait ainsi la conception générale de *fonction automorphe*, qui depuis Poincaré n'avait été considérée pour des fonctions de plus d'une variable que dans des cas très particuliers. D'ailleurs, dans un mémoire ultérieur (1943), Siegel étudia des sous-groupes discrets de $\mathrm{Sp}(2n, \mathbb{R})$ autres que $\mathrm{Sp}(2n, \mathbb{Z})$, opérant de façon proprement discontinue dans D ; il établit le lien entre ces questions et la théorie des groupes de Lie, en observant que le « demi-espace de Siegel » est isomorphe à un domaine borné d'un $\mathbb{C}^N$ qui est un espace symétrique au sens de E. Cartan (complètement déterminés par ce dernier). On a depuis lors développé l'étude des fonctions automorphes correspondant à des groupes opérant de façon proprement discontinue dans ces espaces.

En 1903, P. Epstein avait défini, pour une forme quadratique *définie positive* $Q(x_1, x_2, \ldots, x_n)$ à n variables et à coefficients entiers, des « fonctions zêta » dont la plus simple est :

$$\zeta_Q(s)=\Sigma(Q(x_1, x_2, \ldots, x_n))^{-s}, \tag{35}$$

où la sommation est faite dans $\mathbf{Z}^n-\{0\}$; la série converge pour $\mathrm{Re}\, s>n/2$, et Epstein avait montré qu'elle se prolonge en une fonction méromorphe dans tout le plan, satisfaisant à une équation fonctionnelle du même type que les fonctions zêta des corps de nombres. Cette définition ne pouvait manifestement s'étendre telle quelle aux formes quadratiques indéfinies. Dans deux mémoires publiés en 1938-1939, Siegel montra qu'on parvenait à une telle extension de la façon suivante. Soit S la matrice symétrique d'une forme quadratique de signature $(n, m-n)$, et soit $\Gamma(S)$ le groupe de ses « unités » ; ce groupe opère proprement sur la partie ouverte U de la grassmannienne $G_{m,n}$ formée des sous-espaces vectoriels de dimension n de $\mathbf{R}^m$, dans lesquels la forme quadratique est positive non dégénérée ; en outre, il y a dans U un domaine fondamental de volume fini $\mu(S)$ pour cette opération. Pour un vecteur $a\in\mathbf{Z}^m$, soit de même $\Gamma(S, a)$ le sous-groupe de $\Gamma(S)$ laissant fixe a ; pour $m\geqq 3$, $\Gamma(S, a)$ a encore dans U un domaine fondamental de volume fini $\mu(S, a)$; pour tout entier $t>0$ tel que l'équation ${}^t a . S . a=t$ ait au moins une solution $a\in\mathbf{Z}^m$, Siegel pose :

$$M(S, t)=\sum_a \mu(S, a),$$

la somme étant étendue à un ensemble de solutions de ${}^t a . S . a=t$ représentant les orbites de $\Gamma(S, a)$ dans l'ensemble de toutes les solutions de cette équation dans $\mathbf{Z}^m$. La définition de la fonction zêta est alors :

$$\zeta(S, s)=\sum_{t>0} M(S, t)\, t^{-s}, \tag{36}$$

et Siegel prouve que cette série converge pour $\mathrm{Re}\, s>m/2$, et se prolonge en une fonction méromorphe dans tout le plan et satisfaisant à une équation fonctionnelle. Il obtient cette dernière en la ramenant (suivant l'idée de Riemann) à une identité relative à une fonction thêta (d'une variable cette fois) dépendant d'un paramètre qui parcourt un domaine fondamental de $\Gamma(S)$ dans l'ouvert V de la grassmannienne $G_{m,n}$, identité du même type que (28).

Il devait revenir en 1951-1952 à cette fonction, en obtenant des formules de transformation de cette fonction thêta pour les transformations du groupe modulaire, ainsi qu'une expression de sa « valeur moyenne » dans le domaine fondamental de $\Gamma(S)$

dans U; en outre, de cette « formule de Siegel », il déduisit une autre preuve de son résultat fondamental de 1936 pour les formes quadratiques indéfinies.

C'est aussi en 1951-1952 que Siegel envisage l'extension de ses résultats à des formes quadratiques dont les coefficients et les variables appartiennent à une algèbre simple sur le corps Q, admettant une involution; il indique sans démonstration comment se généraliserait dans ce cadre plus étendu la « formule de Siegel ».

Enfin, il est clair d'après ce qui précède que de plus en plus les méthodes de Siegel reposaient sur l'évaluation de « volumes » de domaines fondamentaux (ou, ce qui revient au même, d'espaces homogènes quotients de groupes de Lie par des sous-groupes discrets); aussi, dès 1936, il consacra plusieurs articles à des calculs explicites de tels volumes.

Toutes ces remarquables techniques forgées par Siegel ont peu après conduit à une vue d'ensemble, où les idées de « mesure » dans les groupes de Lie ou les groupes p-adiques ont abouti à la découverte par Tamagawa d'une mesure privilégiée sur un groupe « adélisé » d'un groupe linéaire algébrique défini sur un corps de nombres algébriques, et du fait que les propriétés de cette mesure entraînaient les théorèmes de Siegel de 1935-1937 sur les formes quadratiques. A. Weil a ensuite interprété de même la « formule de Siegel » de 1951-1952 en théorie adélique.

Dans une toute autre direction, Siegel, dans son mémoire de 1935, avait déduit de sa formule (25) le fait remarquable que la valeur de la fonction zêta d'un corps de nombres en un entier <0 est un nombre rationnel. Il revint beaucoup plus tard à cette question, obtenant des résultats plus précis en se servant des formes modulaires. Ses résultats ont été à la base des progrès récents sur ce sujet, qui fait actuellement l'objet de recherches nombreuses.

Nous avons essayé de cerner les principales directions de l'œuvre imposante de Siegel en Théorie des nombres. Mais il ne s'y est pas limité, et n'a cessé de contribuer par des articles de moindre importance (mais dont aucun n'est banal) à de nombreuses questions d'arithmétique, d'algèbre et d'analyse : géométrie des nombres, nombres de Pisot, valeurs moyennes de fonctions arithmétiques, sommes de carrés et problème de Waring dans les corps de nombres, zéros des fonctions L de Dirichlet, itération des fonctions holomorphes, fonctions méromorphes sur une variété kählérienne compacte, groupes de déplacements non euclidiens, groupes discontinus, fonctions abéliennes, équations différentielles sur le tore, calcul des variations. Son sujet de prédilection (en dehors de la Théorie des nombres), auquel il n'a pas consacré moins de 8 articles s'échelonnant entre 1936 et 1971, fut la Mécanique céleste et les équations différentielles analytiques, notamment les systèmes hamiltoniens; il a raconté dans ses souvenirs sur Frobenius qu'en 1915 il s'imaginait que la Mécanique céleste était la seule science qui ne serait jamais utilisée à des fins guerrières!

Si Siegel n'a pas eu beaucoup d'élèves poursuivant ses recherches, il a cependant toujours aimé l'enseignement, même élémentaire, et a publié de nombreux ouvrages didactiques de Théorie des nombres, de Mécanique céleste et de théorie des fonctions de plusieurs variables complexes; son cours de 1948-1949 à l'Institute for advanced study sur cette dernière théorie a longtemps été le seul texte « moderne » à la disposition des chercheurs.

La valeur exceptionnelle de Siegel a très tôt été reconnue; il a reçu de nombreux doctorats « honoris causa » et était membre associé d'une dizaine d'Académies; notre compagnie l'avait élu correspondant en 1956, et associé étranger en 1973. Il était membre de l'ordre « Pour le Mérite » de la R.F.A. ; en 1978, il a reçu le prix international Wolf (Israël) en même temps que I. M. Gelfand ; c'était la première fois que ce prix était décerné.

Comme le dit un des éditeurs de ses œuvres, « Siegel représente le type idéal de la pensée mathématique contemporaine. A la fois classique et moderne, son œuvre a profondément influencé la culture mathématique de notre temps. »