NANO

Nano

101. — Benoît Kloeckner. *Un bref aperçu de la géométrie projective*
102. — Michel Balazard. *Le théorème des nombres premiers*
103. — Bruno Kahn. *Fonctions zêta et L de variétés et de motifs*
104. — Patrick Dehornoy. *Le calcul des tresses*
105. — Alain Debreil, Rached Mneimné. *Le groupe symétrique* $\mathfrak{S}_4$ *et ses métamorphoses*
106. — Roger Mansuy. *Introduction aux graphes aléatoires (et à la méthode probabiliste)*
107. — Jean-Denis Eiden. *Espaces euclidiens, avec une ouverture vers les espaces préhilbertiens réels*
108. — François Berteloot. *Les familles normales*
109. — Anna Cadoret. *Classification des modules sur les anneaux et représentations linéaires des groupes finis*
110. — Antoine Chambert-Loir. *Théorie de l'information.* Trois théorèmes de Claude Shannon
111. — François Rouvière. *Les transformations de Radon* (en préparation)
112. — Bernard Candelpergher. *Les séries d'Euler à Ramanujan* (en préparation)

Antoine Chambert-Loir

Théorie de l'information

Trois théorèmes de Claude Shannon

Calvage & Mounet

ANTOINE CHAMBERT-LOIR est ancien élève de l'École normale supérieure et agrégé de l'Université. Ses recherches sont principalement motivées par des questions de théorie des nombres, souvent formulées dans le langage de la géométrie algébrique. Avec Johannes Nicaise et Julien Sebag, il a publié un ouvrage de recherche sur l'« intégration motivique » (*Motivic Integration,* Birkhäuser, 2018). Il a aussi publié des ouvrages d'enseignement, dont un d'algèbre (essentiellement) commutative, *(Mostly) Commutative Algebra,* Springer, 2021.
Il est actuellement professeur de mathématiques à l'université Paris Cité, site des Grands-Moulins.

antoine.chambert-loir@u-paris.fr

Mathematics Subject Classification (2020):

94-01 Introductory exposition (textbooks, tutorial papers, etc.) pertaining to information and communication theory
94-A17 Measures of information, entropy
94-A20 Sampling theory in information and communication theory
94-A24 Coding theorems (Shannon theory)
60-J10 Markov chains (discrete-time Markov processes on discrete state spaces)
42-A16 Fourier coefficients, Fourier series of functions with special properties, special Fourier series
42-A20 Convergence and absolute convergence of Fourier and trigonometric series
60-F05 Central limit and other weak theorems

ISBN 978-2-493230-00-3

9 782493 230003

∞ Imprimé sur papier permanent

J'avais imaginé l'appeler *information*, mais le mot était déjà trop utilisé, donc j'ai décidé de l'appeler *incertitude*. (...) Von Neumann m'a dit, « Vous devriez l'appeler entropie, pour deux raisons. D'abord, votre fonction d'incertitude a été utilisée en mécanique statistique sous ce nom ; elle a donc déjà un nom. D'autre part, et c'est plus important, personne ne sait réellement ce qu'est l'entropie, de sorte que dans une discussion, vous aurez toujours l'avantage. »

Claude Shannon (1916–2001),
cité par Tribus & McIrvinc (1971)

À toutes et tous les contributeurs de Wikipedia

Table des matières

III. Échantillonnage

IV. Solutions des exercices

Avant-propos

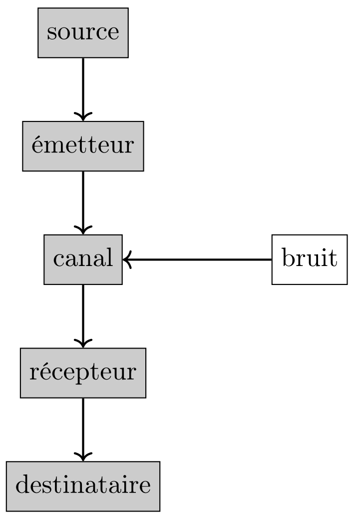
La *théorie mathématique de la communication* vise à étudier de façon mathématiques dans quelles conditions on peut transmettre des données, en particulier à quelle vitesse, et avec quelle fiabilité. C'est à Claude Shannon (1916–2001) que l'on doit cette expression dont une variante a donné naissance à un domaine de recherche, la *théorie de l'information*. Mathématicien, ingénieur électrique, informaticien, cryptologue, l'activité scientifique de Shannon engloba tous ces champs, que ce fût dans le cadre de sa thèse au MIT, de l'armée américaine pendant la Seconde Guerre mondiale, ou du MIT dont il devint professeur en 1956.

On peut trouver plusieurs sources d'inspiration aux travaux de Shannon, comme les travaux des mathématiciens Norbert Wiener et Andrey Kolmogorov, mais ses motivations directes semblent toutes liées à l'effort militaire pendant la Seconde Guerre mondiale. Dans un rapport du Service de recherche de la défense nationale américaine, consacré au contrôle automatique d'armes à feu, Blackman *et al.* (1946) observent

> « l'analogie évidente entre le problème de lisser les données pour éliminer ou réduire l'effet des erreurs de mesure et le problème de séparer un signal du bruit qui le perturbe dans des systèmes de communications. »

Aussi propose-t-il, dans un article écrit en 1945 et publié quelques années plus tard (Shannon, 1949*b*), une analyse théorique des systèmes cryptographiques et pose la question de leur sécurité en termes de théorie de l'information : si l'adversaire dispose d'une partie donnée du message chiffré, de quelle quantité d'information dispose-t-il ?

Après la guerre, Shannon élargit (ou plutôt *efface*) le contexte de ses idées et propose, dans son article fondateur de Shannon (1948),[1] qu'un système de communication soit modélisé par le diagramme suivant (en général disposé en ligne, mais la largeur des pages de ce livre rendait la figure peu lisible) :



La *source* est l'entité qui possède une information, un *message*, à transmettre à son *destinataire*. La source peut être une station de radio ou de télévision, un journal, un site web, vous ou moi désirant annoncer une mauvaise nouvelle au téléphone ou par courrier électronique, une sonde spatiale prenant des photos des planètes qu'elle survole, etc. Le message peut être un texte, une photographie, un morceau de musique, une combinaison de ceux-ci.

[1] Shannon publia cet article l'année suivante sous la forme d'un livre, précédé d'une introduction, et aujourd'hui disponible dans une traduction française (Shannon & Weaver, 2018). On notera que le titre a légèrement changé, passant d'*une* à *la* théorie mathématique de la communication.

L'*émetteur* est l'appareil physique par lequel nous allons diffuser ce message, c'est par exemple un émetteur de radio ou de télévision. À l'époque de Shannon, la transmission était souvent analogique. Dans le cas du téléphone ou de la radio, le son est représenté par l'amplitude de la pression sur le micro que ce dernier transforme en un signal électrique proportionnel. Il s'agit donc de transmettre cette amplitude, représentée par une fonction du temps, ou par deux telles fonctions pour un signal stéréo. Dans le cas de la télévision couleur, il s'agira de transmettre trois amplitudes pour la couleur (rouge/vert/bleu) en chaque point de l'écran, deux amplitudes pour le son, le tout dépendant du temps. De nos jours, la transmission des signaux — télévision, téléphone, Internet — est essentiellement numérique, à l'exception notable de la radio FM, la radio numérique terrestre (DAB) qui peine à décoller : le message est transformé en une suite de nombres qu'il s'agit de transmettre.

Le *récepteur* est l'appareil par lequel le destinataire reçoit ce message, c'est un poste de radio ou de télévision, éventuellement associé à un « décodeur » dans le cas de la télévision numérique terrestre ou de la télévision par Internet, un téléphone, un ordinateur relié au réseau Internet, etc.

Le *canal* est le medium physique par lequel l'information est transmise de l'émetteur au récepteur : l'air pour la transmission de la radio/télévision par voie hertzienne, la fibre optique du fournisseur Internet, les câbles en cuivre du réseau de téléphone, etc. Comme tout objet physique, ce canal est sujet à des perturbations — du *bruit* — par lesquelles le signal qui parvient au récepteur diffère de celui envoyé par l'émetteur, de sorte que le message reçu par le destinataire diffère de celui envoyé par la source.

La théorie mathématique de la communication vise à analyser dans quelles conditions un canal donné, soumis à un certain bruit, peut, ou pas, transmettre un message donné. Deux théorèmes de SHANNON (1948) répondent ainsi aux questions suivantes.

1. À quelle vitesse est-il possible de transmettre un message ?
2. En présence de bruit, est-il possible de transmettre un message de manière fiable ?

Si le canal permet de diffuser c symboles par unité de temps, il semble évident que l'on peut transmettre un message de N symboles pendant un temps N/c, mais peut-on faire mieux ? Ensuite, comment détecter une mauvaise transmission de certains symboles et, éventuellement, les corriger ?

Le présupposé de base de la théorie est que les messages à transmettre, du fait même de leur origine, ne sont pas arbitraires. Si c'est un texte, certaines lettres seront plus fréquentes que d'autres ; si c'est l'enregistrement d'une voix ou d'un morceau de musique certaines fréquences seront absentes du signal, sans même tenir compte du fait que les hauts-parleurs ne peuvent pas les restituer ou l'oreille humaine les percevoir.

Dans son article, SHANNON (1948) propose une définition de la « quantité d'information » contenue dans un message, qu'il appelle *entropie*. Plus exactement, il s'agit de la quantité d'information contenue dans l'ensemble des signaux susceptibles d'être transmis. Sa définition est de nature probabiliste et son étude fait l'objet du chapitre I. Nous y définissons l'entropie d'une variable aléatoire et plusieurs variantes :

- ▷ l'entropie conditionnelle, qui représente l'information supplémentaire qu'offre une variable aléatoire par rapport à une autre ;
- ▷ l'information mutuelle entre deux variables aléatoires, qui représente de façon symétrique l'information que chacune dit de l'autre ;
- ▷ le taux d'entropie d'une suite de variables aléatoires (un *processus stochastique*), qui représente l'information moyenne apportée par chacune d'entre elles.

En particulier, nous calculons le taux d'entropie dans le cas important des *processus markoviens*, dans lesquels chaque variable aléatoire est indépendante de l'ensemble de celles qui précèdent.

Comme ce chapitre est assez théorique, nous l'avons fait précéder d'un chapitre de « rappels » de théorie de probabilités.

Le chapitre II est consacré au *codage*, c'est-à-dire aux deux théorèmes de Shannon évoqués plus haut qui permettent d'analyser deux aspects de la transmission d'un signal : la possibilité de la compression, et la possibilité de corriger les erreurs. Ces deux faits sont à la base de toute la théorie moderne des télécommunications numériques.

En fin de volume, nous discutons la question de l'*échantillonnage* : c'est la première phase de la numérisation d'un signal qui le voit transformé en une suite de nombres. Le « théorème d'échantillonnage », classiquement attribué à Shannon, mais que ce dernier présente dans son article (SHANNON, 1949*a*) comme bien connu, fournit des hypothèses sous lesquelles cet échantillonnage peut se faire sans perte d'information. L'outil principal de ce chapitre est la théorie de Fourier, qui permet de séparer les différentes fréquences qui apparaissent dans un signal, et une partie importante de ce chapitre est consacrée à établir les principaux résultats de la théorie des séries de Fourier et de la transformation de Fourier.

Chaque chapitre se clôt par un bon nombre d'exercices, dont les solutions sont données dans un dernier chapitre.

J'ai ajouté quelques indications historiques pour éclairer les noms propres des mathématiciens ou physiciens que l'on rencontre au cours du livre — trop peu de mathématiciennes hélas. Pas aussi rigoureuses qu'il conviendrait, ces courtes notices sont une invitation à engager une lecture de ces travaux parfois anciens qui ne néglige pas les conditions historiques, sociales, scientifiques, de leur réalisation, qu'il s'agisse par exemple de Joseph Fourier, professeur à l'École polytechnique, recruté par Napoléon au sein de l'équipe scientifique qui accompagna la campagne d'Égypte, puis nommé préfet des Alpes, ou de Claude Shannon dont les travaux, pendant la Seconde Guerre mondiale, visèrent en partie à résister aux efforts de brouillage des télécommunications électriques.

Rédiger ces notices m'a fait (re)prendre conscience de la façon dont les sciences mathématiques et physiques, la théorie de l'information en particulier, étaient intimement liées à l'effort militaire, plus généralement au complexe militaro-industriel. À un moment de l'histoire où les mathématiques et l'informatique jouent un grand rôle via le traitement d'énormes quantités de données numériques ou le chiffrement des télécommunications, au point que l'on puisse parler de « capitalisme de surveillance », que cette conscience touche les lecteurs et lectrices de ce livre, et ce maigre effort n'aura pas été absolument vain !

La littérature anglophone propose beaucoup d'ouvrages d'introduction à la théorie de l'information. Écrits par des mathématiciens plus compétents que moi en ce domaine, ces livres sont souvent excellents, plus complets, et vont parfois beaucoup plus loin. En particulier, le livre de Cover & Thomas (2006) m'a été très utile, tant pour la présentation générale que pour de nombreux exercices, et les lecteurices qui sauront en profiter reconnaîtront ma dette à son égard. J'espère que le présent livre, de taille et d'ambition limitée, pourra servir d'introduction aimable à ces belles questions.

J'ai enseigné ce cours pendant trois années à l'université de Paris (ex-Diderot), au sein du Master 1 maths-info. Je remercie Georges Skandalis et Justin Salez de m'avoir confié les notes et des énoncés d'exercices qu'ils utilisaient pour ce cours. Je remercie aussi Guillaume Garrigos pour avoir assuré diligemment les séances de travaux dirigés et pour avoir insisté que les arguments que je donnais méritaient souvent de plus amples détails.

Je remercie aussi Laurant Dauday, Pierre Karpman, Roger Mansuy et Pascal Molin des commentaires qu'ils m'ont envoyés.

Rached Mneimné a accueilli chaleureusement ce petit ouvrage dans sa collection. Avec Alain Debreil, il a relu l'ensemble de façon pointilleuse ; Alain s'est en outre chargé d'ajouter quelques diagrammes. Merci à eux !

Je remercie enfin les étudiantes et les étudiants du master pour leur participation, en particulier au cours de l'année universitaire 2020–2021 pendant laquelle la pandémie de CoVid-19 les a contraint d'étudier à distance.

Chapitre 0

Éléments de théorie des probabilités

Ce chapitre a pour but de rappeler quelques définitions et résultats élémentaires en théorie des probabilités discrètes.

On doit à KOLMOGOROV (1956)[1] (la première édition, en allemand, date de 1933) d'avoir fondé la théorie des probabilités sur la base de l'intégrale de Lebesgue (voir note en bas de la page 130). Les variables aléatoires discrètes que nous considérons dans tout ce petit livre ne nécessitent pas un tel bagage et, dans la plupart des cas, les calculs résulteront de manipulations de sommes finies. Parfois, les variables aléatoires discrètes peuvent prendre une infinité de valeurs et les sommes finies doivent alors être remplacées par des séries ou plutôt par des familles sommables, car leur ensemble d'indices n'est pas naturellement bien ordonné. Le premier paragraphe rappelle cette théorie ; il est peut-être malin de ne pas y prêter une trop grande attention.

[1]Andrey KOLMOGOROV (1903–1987), mathématicien soviétique dont les travaux fondamentaux concernèrent la théorie des probabilités, la turbulence en mécanique des fluides, la mécanique céleste, la théorie de la complexité, les statistiques, la topologie...

Nous donnons ensuite une définition formelle, un peu trop formelle même, d'un espace probabilisé et d'une variable aléatoire discrète sur un tel espace. Pour nous éviter des contorsions liées à d'éventuels événements de probabilité nulle, j'ai été conduit à supposer l'univers « complet » pour la probabilité considérée.

La section suivante rappelle les définitions de variables aléatoires (discrètes), de leur loi, leur espérance et leur variance (quand elles existent).

Nous rappelons enfin la notion d'indépendance, cruciale en théorie des probabilités. Elle est relativement élémentaire en ce qui concerne l'indépendance de deux événements, un peu plus subtile lorsqu'il s'agit d'indépendance de deux variables aléatoires. On termine cette section par les notions d'espérance et de variance conditionnelle.

1. Familles sommables

On va être amené à manipuler des sommes infinies, indexées par un ensemble qui n'est pas forcément l'ensemble des entiers, ni en bijection évidente avec l'ensemble des entiers. La théorie des familles sommables a pour but de préciser dans quel contexte ces sommes existent et de les calculer.

1.1. Définition. Soit A un ensemble, et soit $(z_a)_{a\in A}$ une famille de nombres complexes indexée par l'ensemble A. On dit que cette famille est *sommable* s'il existe un nombre complexe z tel que, pour tout $\varepsilon > 0$, il existe une partie finie B de A telle que

$$\Big|z - \sum_{a\in C} z_a\Big| \leqslant \varepsilon,$$

pour toute partie finie C de A telle que $B \subseteq C$. On dit alors que (z_a) est sommable de somme z.

Si l'ensemble A est fini, la famille $(z_a)_{a\in A}$ est sommable de somme la somme $z = \sum_{a\in A}$ de cette famille finie : dans la définition, il suffit de prendre $B = A$.

1.2. Lemme. *Soit (z_a) une famille et soient z, z' des nombres complexes. Supposons que la famille (z_a) soit sommable de somme z et soit sommable de somme z'. Alors, $z = z'$.*

Compte tenu de ce lemme, il est légitime d'appeler *somme* d'une famille sommable (z_a) l'unique nombre complexe z tel qu'elle soit sommable de somme z ; on le note $\sum_{a\in A} z_a$.

Toute la théorie des familles sommables vise à donner des moyens de calculer cette somme, en particulier de justifier sous quelles conditions elle se comporte « comme » une somme finie.

Démonstration. Soit ε un nombre réel > 0. Soit B une partie finie de A telle que $|z - \sum_{a\in C} z_a| \leqslant \varepsilon$ pour toute partie finie C contenant B. Choisissons B' de façon analogue pour z' et posons $C = B \cup B'$. Alors, $|z - \sum_{a\in C} z_a| \leqslant \varepsilon$ et $|z' - \sum_{a\in C} z_a| \leqslant \varepsilon$, de sorte que $|z - z'| \leqslant 2\varepsilon$. Comme ε est arbitraire, cela entraîne $z = z'$. $\square$

1.3. Lemme. *Pour qu'une famille $(z_a)_{a\in A}$ de nombres réels positifs soit sommable, il faut et il suffit que les sommes $\sum_{a\in B} z_a$ soient majorées, lorsque B parcourt l'ensemble des parties finies de A. Alors,*

$$\sum_{a\in A} z_a = \sup_{\substack{B\subseteq A \\ B \text{ fini}}} \Big(\sum_{a\in B} z_a\Big).$$

Lorsque ces sommes ne sont pas majorées, on pose $\sum_{a\in A} z_a = +\infty$.

Démonstration. Considérons l'ensemble des sommes $\sum_{a\in B} z_a$, où B parcourt l'ensemble des parties finies de A, et notons z la borne supérieure de cet ensemble. C'est un nombre réel si ces sommes sont majorées, et $+\infty$ sinon. Soit $\varepsilon > 0$, et soit B une partie finie de A telle que $z - \varepsilon < \sum_{a\in B} z_a \leqslant z$. Soit C une partie finie de A telle que $B \subseteq C$. On a $\sum_{a\in C} z_a \leqslant z$, par définition de z. Par ailleurs, comme $z_a \geqslant 0$ pour tout $a \in A$, on a

$$\sum_{a\in C} z_a = \sum_{b\in B} z_b + \sum_{a\in C - B} z_a \geqslant \sum_{b\in B} z_b \geqslant z - \varepsilon.$$

Par conséquent, $|z - \sum_{a\in C} z_a| \leqslant \varepsilon$. Cela démontre que la famille $(z_a)_{a\in A}$ est sommable, de somme z. $\square$

1.4. Donnons quelques propriétés des familles sommables.

1. *Si les familles* $(z_a)_{a\in A}$ *et* $(z'_a)_{a\in A}$ *sont sommables, il en est de même de la famille* $(z_a + z'_a)_{a\in A}$*, et l'on a*
$$\sum_{a\in A}(z_a + z'_a) = \sum_{a\in A} z_a + \sum_{a\in A} z'_a.$$
Posons $z = \sum_{a\in A} z_a$ et $z' = \sum_{a\in A} z'_a$. Soit $\varepsilon > 0$, et soit B une partie finie de A telle que $|z - \sum_{a\in C} z_a| < \varepsilon/2$ pour toute partie finie C de A telle que $B \subseteq C$. De même, soit B' une partie finie de A telle que $|z' - \sum_{a\in C} z'_a| < \varepsilon/2$ pour toute partie finie C de A telle que $B' \subseteq C$. L'ensemble $B \cup B'$ est fini ; de plus, pour toute partie finie C de A telle que $B \cup B' \subseteq C$, on a
$$\begin{aligned}\Big|(z+z') - \sum_{a\in C}(z_a + z'_a)\Big| &\leqslant \Big|z - \sum_{a\in C} z_a\Big| + \Big|z' - \sum_{a\in C} z'_a\Big| \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon.\end{aligned}$$
Cela démontre que la famille $(z_a + z'_a)_{a\in A}$ est sommable et que sa somme est égale à $z + z'$.

2. *Soit* λ *un nombre complexe. Si la famille* $(z_a)_{a\in A}$ *est sommable, il en est de même de la famille* $(\lambda z_a)_{a\in A}$*, et l'on a*
$$\sum_{a\in A}\lambda z_a = \lambda\sum_{a\in A} z_a.$$
Soit $\varepsilon > 0$ et soit $z = \sum_{a\in A} z_a$. Soit B une partie finie de A telle que pour toute partie finie C de A telle que $B \subseteq C$, on ait $|z - \sum_{a\in C} z_a| < \varepsilon/(1+|\lambda|)$. Alors, pour toute telle partie C, on a
$$\Big|\lambda z - \sum_{a\in C}\lambda z_a\Big| = |\lambda|\times\Big|z - \sum_{a\in C} z_a\Big| < \varepsilon\frac{|\lambda|}{1+|\lambda|} < \varepsilon,$$
ce qui démontre que la famille $(\lambda z_a)_{a\in A}$ est sommable et que sa somme est égale à λz.

3. *Si la famille* $(z_a)_{a\in A}$ *est sommable, il en est de même de la famille* $(\overline{z_a})_{a\in A}$*, et l'on a*
$$\sum_{a\in A}\overline{z_a} = \overline{\sum_{a\in A} z_a}.$$

Soit $\varepsilon > 0$ et soit $z = \sum_{a\in A} z_a$. Soit B une partie finie de A telle que pour toute partie finie C de A telle que $B \subseteq C$, on ait $|z - \sum_{a\in C} z_a| < \varepsilon$. Alors, pour toute telle partie C, on a

$$\left|\overline{z} - \sum_{a\in C} \overline{z_a}\right| = \left|z - \sum_{a\in C} z_a\right| < \varepsilon,$$

ce qui démontre que la famille $(\overline{z_a})_{a\in A}$ est sommable et que sa somme est égale à $\overline{z}$.

1.5. Proposition

1. (Critère de Cauchy[2])
   Pour qu'une famille $(z_a)_{a\in A}$ soit sommable, il faut et il suffit que pour tout $\varepsilon > 0$, il existe une partie finie B de A telle que l'on ait $|\sum_{a\in C} z_a| \leqslant \varepsilon$ pour toute partie finie C de A qui est disjointe de B.
2. *Pour qu'une famille $(z_a)_{a\in A}$ soit sommable, il faut et il suffit qu'il existe un nombre réel M tel que $|\sum_{a\in B} z_a| \leqslant M$ pour toute partie finie B de A.*

Démonstration

1. La nécessité de ce critère est tout à fait semblable à celle du critère de Cauchy pour les séries. Supposons que la famille (z_a) soit sommable de somme z. Soit $\varepsilon > 0$; choisissons une partie finie B de A telle que

$$\left|z - \sum_{a\in C} z_a\right| \leqslant \varepsilon/2$$

pour toute partie finie C de A contenant B.

[2]Augustin Louis Cauchy (1789–1857) était un mathématicien français, dont l'œuvre a été très importante, en analyse réelle et complexe, ainsi qu'en algèbre. Bien que ses arguments aient manqué de rigueur, on lui doit les premières définitions précises de convergence de séries, de limites, de dérivabilité, le théorème des valeurs intermédiaires. Il a aussi travaillé en mécanique céleste. Sa négligence à l'encontre des mémoires qu'Abel et Galois avaient soumis à l'Académie des sciences a certainement retardé la diffusion des découvertes de ces derniers.

Soit alors une partie finie C de A qui est disjointe de B ; on a

$$\begin{aligned}\Big|\sum_{a\in C} z_a\Big| &= \Big|\sum_{a\in B\,\cup\, C} z_a - \sum_{a\in B} z_a\Big| \\ &\leqslant \Big|\sum_{a\in B\,\cup\, C} z_a - z\Big| + \Big|\sum_{a\in B} z_a - z\Big| \\ &\leqslant 2\varepsilon/2 = \varepsilon.\end{aligned}$$

Pour démontrer que cette condition est suffisante, définissons par récurrence une suite croissante (B_n) de parties finies de A en posant $B_0 = \emptyset$, puis, si $n \geqslant 1$ est tel que B_{n-1} est définie, prenons pour partie B_n une partie finie de A contenant B_{n-1} telle que l'on ait $|\sum_{a\in C} z_a| \leqslant 1/n$ pour toute partie finie C de A qui est disjointe de B_n. Posons alors, pour tout entier n, $u_n = \sum_{a\in B_n} z_a$. Si m et n sont des entiers tels que $n \geqslant m \geqslant 1$, et l'on a

$$|u_m - u_n| = \Big|\sum_{a\in B_n} z_a - \sum_{a\in B_m} z_a\Big| = \Big|\sum_{a\in B_n - B_m} z_a\Big| \leqslant \frac{1}{m}.$$

Ainsi, la suite (u_n) de nombres complexes vérifie le critère de Cauchy, donc converge dans $\mathbf{C}$; notons u sa limite. Faisant tendre n vers $+\infty$, on en déduit l'inégalité $|u_m - u| \leqslant 1/m$ pour tout entier $m \geqslant 1$.

Démontrons que la famille (z_a) est sommable de somme z. Soit $\varepsilon > 0$, et soit m un entier tel que $\dfrac{2}{m} < \varepsilon$. Soit C une partie finie de A contenant B_m ; on a

$$\begin{aligned}\Big|u - \sum_{a\in C} z_a\Big| &= \Big|(u-u_m) - \sum_{a\in C - B_m} z_a\Big| \\ &\leqslant |u-u_m| + \Big|\sum_{a\in C - B_m} z_a\Big| \\ &\leqslant \frac{1}{m} + \frac{1}{m} \leqslant \varepsilon.\end{aligned}$$

Cela prouve le résultat voulu.

2. La condition est nécessaire. Supposons en effet que la famille $(z_a)_{a\in A}$ soit sommable, de somme z, et soit C_1 une partie