

Comment démontrer des formules sans effort ?

EXPOSÉ DE MAÎTRISE

Marc MEZZAROBBA Sam ZOGHAIB

Sujet proposé par François LOESER

Résumé

Nous exposons un ensemble de méthodes qui permettent d'évaluer « en forme close », c'est-à-dire sans signe \sum , de vastes classes de sommes discrètes, et de trouver des démonstrations faciles à vérifier des identités obtenues. Les principaux ingrédients sont l'algorithme de Gosper, qui traite le cas des sommes « indéfinies » $\sum^n t(k)$, celui de Zeilberger, qui « devine » une relation de récurrence vérifiée par une somme « définie » $\sum_k t(n, k)$, celui de Petkovšek, qui résout ce type de récurrence, et les certificats WZ, une façon originale de démontrer des formules.

Ces méthodes forment le cœur de l'ouvrage de Petkovšek, Wilf et Zeilberger, $A = B$ [1]. Nous ne prétendons pas rivaliser en pédagogie avec ses auteurs, aussi nous essayons plutôt de donner une présentation plus concise des seuls résultats importants, sans pour autant omettre de justification.

Table des matières

Introduction	3
1 Termes et récurrences	3
1.1 Quelques notations	3
1.2 Opérateurs sur les suites	4
1.3 Termes hypergéométriques	4
1.3.1 Définition	4
1.3.2 Termes associés	5
1.4 Forme normale des fractions rationnelles	6
2 Sommation	7
2.1 Sommation indéfinie : algorithme de Gosper	8
2.1.1 Équation de Gosper	8
2.1.2 Degré des solutions	9
2.1.3 Solutions en forme close	10
2.2 Sommation définie : algorithme de Zeilberger	10
2.2.1 Termes hypergéométriques propres	11
2.2.2 Algorithme de Zeilberger	14
3 Formules closes et preuves	15
3.1 Preuves WZ	16
3.2 Résolution des récurrences : algorithme de Petkovšek	17
3.2.1 Calcul d'une solution hypergéométrique	18
3.2.2 Espace des solutions	20
Conclusion	22
Références	22

Introduction

Cet exposé porte sur certaines identités entre fonctions discrètes, de celles qui apparaissent naturellement dans les manipulations « combinatoires ». Un exemple typique est l'identité de Dixon

$$\sum_k (-1)^k \binom{a+b}{a+k} \binom{b+c}{b+k} \binom{c+a}{c+k} = \frac{(a+b+c)!}{a!b!c!}.$$

On s'intéresse à la manipulation automatique, par un logiciel de calcul formel, de ce genre de somme.

Deux problèmes se posent. Premièrement, face à une somme comme celle du membre gauche, comment la calculer, autrement dit en trouver une expression débarrassée du signe \sum ? Deuxièmement, comment démontrer à moindres frais de telles identités, qu'elles aient été trouvées par l'ordinateur ou conjecturées par un humain? Il est souhaitable d'obtenir des démonstrations autonomes, faciles à vérifier à la main, qui n'exigent de faire confiance ni au logiciel qui a donné la formule, ni à un gros « méta-théorème » dont chaque preuve ne serait qu'une instance.

Pour certifier de façon relativement aisée à vérifier une identité compliquée, une démarche naturelle serait d'exhiber une relation de récurrence satisfaite par les deux membres, et de s'assurer qu'ils coïncident en un nombre convenable de valeurs initiales. C'est en effet une des idées de base qui reviendront tout le long du texte.

Le point de départ de presque toutes nos manipulations est l'algorithme de Gosper, qui permet de trouver à coup sûr, dès qu'elle existe, la « primitive »¹ d'une suite hypergéométrique. Nous présenterons plusieurs autres méthodes, dues notamment à Zeilberger, qui permettent, en utilisant astucieusement des variantes de cet algorithme, de trouver des récurrences, voire des formules closes, pour des sommes définies, ou encore des preuves peu naturelles mais vérifiables en quelques lignes d'un grand nombre d'identités.

1 Termes et récurrences

Dans cette section, on étudie quelques propriétés des suites hypergéométriques à valeurs dans un corps, qui fonderont les manipulations des sections suivantes.

1.1 Quelques notations

Sauf mention contraire, dans tout l'exposé, \mathbb{K} est un corps de caractéristique nulle.

Notation. Soient $x \in \mathbb{K}$ et $k \in \mathbb{N}$. On introduit les deux notations

$$\begin{aligned} x^{\underline{k}} &= x(x-1)\dots(x-k+1) && \text{(factorielle descendante)} \\ x^{\overline{k}} &= (x+1)\dots(x+k) && \text{(factorielle montante).} \end{aligned}$$

¹l'image réciproque par l'opérateur aux différences finies

(Soulignons qu'avec cette définition, et contrairement à une autre convention fréquente, les k facteurs de $x^{\bar{k}}$ ne commencent qu'à $x + 1$.)

Notation. Si p et q sont deux polynômes, $p \wedge q$ désigne leur pgcd normalisé.

Notation. Si p est un polynôme en une variable n , on note $\lambda(p)$ son coefficient dominant et $(p(n))[n^j]$ le coefficient de son monôme en n^j .

On s'autorise aussi certains abus de notation dans la manipulation des polynômes et fractions rationnelles, considérés tantôt comme tels, tantôt comme fonctions, sans distinction entre variables et indéterminées. Deux polynômes seront donc vus comme égaux s'il coïncident en les valeurs considérées de leurs variables.

1.2 Opérateurs sur les suites

Soit \mathbb{K} un corps. On introduit sur $\mathbb{K}^{\mathbb{N} \times \mathbb{Z}}$ les opérateurs linéaires de décalage de chacune des variables

$$N : t(n, k) \mapsto t(n + 1, k) \quad \text{et} \quad K : t(n, k) \mapsto t(n, k + 1).$$

Indifféremment sur $\mathbb{K}^{\mathbb{N}}$ ou $\mathbb{K}^{\mathbb{Z}}$, on dispose de l'opérateur $t(n) \mapsto t(n + 1)$, également noté N . Dans ce cas $\Delta = N - 1$ désigne l'opérateur aux différences finies.

Un de nos principaux objectifs, nous l'avons dit, est de trouver des récurrences vérifiées par des suites données sous forme de sommes. Nous cherchons plus précisément des récurrences linéaires à coefficients polynomiaux.

Définition. Un *opérateur de récurrence linéaire d'ordre J* (à coefficients polynomiaux) est un opérateur de la forme

$$L = \sum_{j=0}^J a_j(n) N^j$$

où les a_j sont des polynômes en n , avec a_0 et a_J non nuls.

Nous tenons là notre premier procédé de démonstration : pour montrer que deux suites sont égales, il suffit de montrer qu'elles satisfont une même récurrence de cette forme, puis de vérifier l'égalité en un nombre fini de valeurs. Ce nombre n'est pas en général l'ordre de la récurrence, puisque les coefficients peuvent s'annuler. Cependant, ils ne s'annulent qu'un nombre fini de fois, de sorte que la récurrence garantit quand même l'égalité des deux suites à partir d'un certain rang.

1.3 Termes hypergéométriques

1.3.1 Définition

On souhaite automatiser des calculs relativement complexes et *a priori* guère systématiques. Une idée récurrente sera de les ramener à des manipulations algorithmiques de routine : algèbre linéaire de base et opérations sur les polynômes.

Il est donc assez naturel de prendre pour objets de base les termes hypergéométriques, qui sont² en quelque sorte les fonctions les plus générales dont la manipulation se réduit facilement à celle de fractions rationnelles, et donc de polynômes.

Définition. Une suite $(t(k))_{k \in \mathbb{Z}}$ d'éléments d'un corps \mathbb{K} est un *terme hypergéométrique* si et seulement si $t_0 \neq 0$ et t_{k+1}/t_k est une fraction rationnelle en k sans pôle entier. Si les a_i et b_j sont respectivement les racines et les pôles de t_{k+1}/t_k dans une clôture algébrique de \mathbb{K} , comptés avec multiplicité, cela équivaut à

$$t(k) = t_0 \frac{a_1^{k-1} \dots a_p^{k-1}}{b_1^{k-1} \dots b_q^{k-1}} x^k.$$

On a des définitions analogues pour les suites indicées par \mathbb{N} ou par n'importe quel intervalle d'entiers, avec les restrictions correspondantes sur les pôles du rapport de deux termes. La notion s'étend naturellement à plusieurs variables; en particulier, $t(n, k)$ sera dit doublement hypergéométrique lorsque $t(n+1, k)/t(n, k)$ et $t(n, k+1)/t(n, k)$ appartiennent à $\mathbb{K}(n, k)$, et $t(0, 0) \neq 0$.

Une *formule close* pour une suite $t(n)$ est une expression explicite de cette suite comme combinaison linéaire de termes hypergéométriques. Une formule close n'est pas nécessairement une formule « simple », et réciproquement. La définition est tout de même raisonnable : elle signifie qu'on considère que l'on a calculé une somme si on a réussi à l'écrire comme somme d'un nombre constant, et fini, de termes.

1.3.2 Termes associés

Définition. On dit que deux termes hypergéométriques s et t sont associés si et seulement si leur rapport est une fraction rationnelle. Cela définit une relation d'équivalence.

Lemme. Une somme de termes hypergéométriques associés est nulle ou elle-même hypergéométrique.

Démonstration. Si $s(n)$ et $t(n) \neq -s(n)$ sont deux termes hypergéométriques associés,

$$\frac{s(n+1) + t(n+1)}{s(n) + t(n)} = \frac{\frac{s(n+1)}{s(n)} \frac{s(n)}{t(n)} + \frac{t(n+1)}{t(n)}}{\frac{s(n)}{t(n)} + 1}$$

est une fraction rationnelle. □

Proposition. Toute combinaison linéaire de termes hypergéométriques s'écrit de façon unique (à l'ordre près) comme somme de termes deux à deux non associés.

Démonstration. D'après le lemme, il suffit de regrouper les termes associés pour obtenir l'écriture voulue.

Pour l'unicité, montrons par récurrence sur $p \geq 1$ qu'une somme de p termes hypergéométriques deux à deux non associés n'est pas nulle. C'est évident si

²Presque : il existe des classes plus vastes de termes pouvant prétendre à ce titre, à commencer par les termes q -hypergéométriques de Heine. Il se trouve que les résultats que nous exposons s'y généralisent assez bien.

$p = 1$ ou 2 . Soit $p \geq 2$, et supposons $\sum_{i=1}^{p+1} t_i = 0$. Alors $\sum_{i=1}^{p+1} t_i(n+1) = 0$, et $\frac{t_{p+1}(n+1)}{t_{p+1}(n)} \sum_{i=1}^{n+1} t_i(n) = 0$. Soustrayons la seconde relation de la première :

$$\sum_{i=1}^p \left(\frac{t_{p+1}(n+1)}{t_{p+1}(n)} - \frac{t_i(n+1)}{t_i(n)} \right) t_i(n) = 0.$$

Si l'un des termes de la somme est nul, c'est que t_{p+1} et un certain t_i , $i \leq p$ sont proportionnels, et donc associés. Sinon, chaque terme est un terme hypergéométrique, et par hypothèse de récurrence il y en a deux d'associés, *i.e.* il existe $i \neq j$ tels que

$$\frac{\frac{t_{p+1}(n+1)}{t_{p+1}(n)} - \frac{t_i(n+1)}{t_i(n)}}{\frac{t_{p+1}(n+1)}{t_{p+1}(n)} - \frac{t_j(n+1)}{t_j(n)}} \frac{t_i(n)}{t_j(n)} \in \mathbb{K}(n).$$

Le premier facteur étant une fraction rationnelle, t_i et t_j sont associés.

Si maintenant $s(n)$ s'écrit de deux façons $t_1(n) + \dots + t_p(n)$ et $t'_1(n) + \dots + t'_q(n)$ comme somme de termes hypergéométriques deux à deux non associés, $t_1(n) + \dots + t_p(n) - t'_1(n) - \dots - t'_q(n) = 0$ ne peut s'écrire comme somme non vide de termes hypergéométriques non associés ; ainsi chacun des t_i est associé, et donc égal, à exactement un des t'_j . \square

Signalons enfin un résultat immédiat mais utile.

Proposition. Soient L un opérateur de récurrence linéaire à coefficients polynomiaux et $t(n) \notin \ker L$ un terme hypergéométrique. Alors $Lt(n)$ est un terme hypergéométrique associé à $t(n)$.

1.4 Forme normale des fractions rationnelles

Gosper a introduit une écriture des fractions rationnelles qui s'avèrera fondamentale à plusieurs reprises dans la suite.

Soit \mathbb{K} un corps de caractéristique nulle. Supposons en outre que l'on dispose d'un algorithme pour, étant donné $P \in \mathbb{K}[X]$, calculer l'ensemble de ses racines entières³.

Théorème 1. Toute fraction rationnelle $R \in \mathbb{K}(n) \setminus \{0\}$ s'écrit sous la forme

$$R(n) = c \frac{p(n)}{q(n+1)} \frac{r(n+1)}{r(n)} \quad (1)$$

où $c \in \mathbb{K}$ et où p, q, r sont des polynômes unitaires vérifiant $p(n) \wedge r(n) = q(n) \wedge r(n) = 1$ et $p(n) \wedge q(n+h) = 1$ pour tout $h \in \mathbb{N}^*$. L'algorithme de la figure 1 calcule c, p, q et r .

³C'est vrai dans tous les cas qui nous intéressent en pratique : si $\mathbb{K} = \mathbb{Q}$, on peut se ramener à $P \in \mathbb{Z}[X]$, diviser P par une puissance de X telle que 0 ne soit plus racine, et les racines entières restantes divisent alors le terme constant. Si $\mathbb{K} = k(\alpha)$ et si l'on sait trouver les racines entières d'un polynôme de $k[X]$, on pose $Q(\alpha)P = \sum_{i=1}^d P_i(X)\alpha^i$ avec $d < \deg_k \alpha$ ($= \infty$ si α est transcendant), et l'on est ramené à chercher les racines entières communes des P_i .

Donnée. Une fraction rationnelle $R \in \mathbb{K}(n)$.

Résultat. Un quadruplet (c, p, q, r) qui vérifie les hypothèses du théorème 1.

- ▶ Soient $c \in \mathbb{K}$, $p_0, q_0 \in \mathbb{K}[n]$ unitaires premiers entre eux, tels que $R(n) = c p_0(n)/q_0(n+1)$.
- ▶ Soit $\rho(h)$ le résultant des polynômes (en n) $p_0(n)$ et $q_0(n+h)$.
- ▶ Soient $h_1 < \dots < h_m$ les racines entières de $\rho(h)$.
- ▶ Soit $r_0(n) = 1$.
- ▶ Pour i de 1 à m
 - ▶ Soit $d_i(n) = p_{i-1}(n) \wedge q_{i-1}(n+h_i)$.
 - ▶ Soit $p_i(n) = p_{i-1}(n)/d_i(n)$.
 - ▶ Soit $q_i(n) = q_{i-1}(n)/d_i(n-h_i)$.
 - ▶ Soit $r_i(n) = d_i(n-1) \dots d_i(n-h_i) r_{i-1}(n)$.
- ▶ Renvoyer $(c, p, q, r) = (c, p_m, q_m, r_m)$.

FIG. 1 – Mise en forme normale.

Démonstration. On a évidemment $p_m \mid p_{m-1} \mid \dots \mid p_0$ et $q_m \mid q_{m-1} \mid \dots \mid q_0$. À chaque itération i de la boucle *Pour*, on retire le facteur commun $d_i(n)$ de $p_0(n)$ et $q_0(n+h)$, en conservant

$$c \frac{p_i(n)}{q_i(n+1)} \frac{r_i(n+1)}{r_i(n)} = c \frac{p_{i-1}(n)}{q_{i-1}(n+1)} \frac{r_{i-1}(n+1) d_i(n) \dots d_i(n-h+1)}{r_{i-1}(n) d_i(n-1) \dots d_i(n-h)} = R(n),$$

donc le résultat de l'algorithme est bien une écriture de $R(n)$. Un entier h tel que $p_{i-1}(n) \wedge q_{i-1}(n+h_i) \neq 1$ est nécessairement la différence d'une racine de q_0 et d'une racine de p_0 , donc vérifie $\rho(h) = 0$. Les h_i étant traités par ordre croissant, $p_i(n)$ et $q_i(n+h)$ sont premiers entre eux pour tout $h \in \llbracket 1, h_{i+1} - 1 \rrbracket$. Tous les facteurs communs sont retirés à une étape ou à une autre, et finalement $q(n+h)$ est premier avec $p(n)$ pour tout $h \in \mathbb{N}^*$.

Supposons $p \wedge r \neq 1$. Par définition de r , il existe i et $j \in \llbracket 1, h_i \rrbracket$ tels que $p(n) \wedge d_i(n-j) \neq 1$. Puisque $p(n) \mid p_i(n)$ et $d_i(n) \mid q_i(n+h_i)$, on en déduit $p_i(n) \wedge q_i(n+h_i-j) \neq 1$, ce qui est impossible. Un raisonnement similaire montre que $q \wedge r = 1$. \square

On peut montrer que la factorisation (1) est en réalité unique, ce qui justifie son nom de forme normale. Parmi les (c, p, q, r) vérifiant seulement (1) et $p(n) \wedge q(n+h) = 1$, la forme normale minimise le degré de r .

2 Sommation

Cette section pourrait s'intituler « de la somme à la récurrence » : il s'agit d'expliquer comment trouver une récurrence vérifiée par une somme, définie ou non. Si la récurrence est du premier ordre, on en déduit facilement une formule close.

2.1 Sommation indéfinie : algorithme de Gosper

Soit $t(k)$ un terme hypergéométrique. On cherche à calculer en forme close la somme $\sum_{k=0}^{n-1} t(k)$ en fonction de n , autrement dit à trouver une somme $T(k)$ d'un nombre *constant* de termes hypergéométriques telle que $T(k+1) - T(k) = t(k)$. Par analogie avec le cas continu, on dira alors que T est une primitive discrète de t .

Gosper a donné en 1978 un algorithme qui répond complètement à cette question : sa méthode trouve une primitive⁴ de $t(k)$ en forme close ou prouve qu'il n'en existe pas. Sommairement, le principe est le suivant : on ramène, par une transformation astucieuse, l'équation $\Delta T = t$ à une équation dont l'inconnue est *polynomiale*. On borne ensuite le degré des solutions, et on résout finalement la nouvelle équation par simple identification des coefficients. Voyons cela plus en détail.

2.1.1 Équation de Gosper

Cherchons les solutions hypergéométriques de

$$T(k+1) - T(k) = t(k).$$

La première étape de l'algorithme de Gosper consiste à écrire $t(k+1)/t(k)$ sous la forme normale présentée à la section 1.4. Soit donc (c, p, q, r) cette forme normale. Rappelons la propriété que l'on va utiliser ici, à savoir que $p(k)$ et $q(k+h)$ sont premiers entre eux pour tout $h \in \mathbb{N}^*$. Changeons d'inconnue en posant

$$T(k) = \frac{q(k)}{r(k)} t(k) s(k).$$

En se rappelant

$$t(k+1) = c \frac{p(k)}{q(k+1)} \frac{r(k+1)}{r(k)} t(k),$$

il vient

$$\begin{aligned} T(k+1) - T(k) &= \frac{q(k+1)}{r(k+1)} t(k+1) s(k+1) - \frac{q(k)}{r(k)} t(k) s(k) \\ &= c \frac{p(k)}{r(k)} t(k) s(k+1) - \frac{q(k)}{r(k)} t(k) s(k). \end{aligned}$$

L'équation $T(k+1) - T(k) = t(k)$ équivaut donc à

$$cp(k) s(k+1) - q(k) s(k) = r(k). \quad (2)$$

L'inconnue $s(k)$ n'est *a priori* qu'un terme hypergéométrique. Mais on a le théorème suivant, qui est le point-clé de la méthode de Gosper.

Théorème 2 (Gosper). Soit $s(k)$ une solution hypergéométrique de (2). Alors $s(k)$ est un polynôme.

⁴Notons en passant que l'analogie de l'algorithme de Gosper pour le problème de l'intégration existe : il s'agit de l'algorithme de Risch, qui permet de calculer, dès qu'il en existe, une primitive élémentaire d'une fonction élémentaire.

Démonstration. Déjà, $T(k)$ étant supposé hypergéométrique,

$$s(k) = \frac{r(k) T(k)}{q(k) t(k)} = \frac{r(k) T(k)}{q(k) T(k+1) - T(k)} = \frac{r(k)}{q(k)} \frac{1}{T(k+1)/T(k) - 1}$$

est une fraction rationnelle. Posons $s(k) = a(k)/b(k)$ avec $a, b \in \mathbb{K}[k]$ premiers entre eux, et supposons par l'absurde b non constant. Soit

$$\mu = \max\{\mu \in \mathbb{N} \mid b(k) \wedge b(k+\mu) \neq 1\}$$

et soit $u(k)$ un facteur commun irréductible de $b(k)$ et $b(k+\mu)$. Alors $u(k+1)$ divise $b(k+1)$, or comme $s(k)$ est solution de (2),

$$cp(k)a(k+1)b(k) - q(k)a(k)b(k+1) = r(k)b(k)b(k+1),$$

donc $u(k+1) \mid p(k)a(k+1)b(k)$. Mais $u(k+1)$ ne divise ni $a(k+1)$ (puisque $a(k) \wedge b(k) = 1$) ni $b(k)$ (sans quoi $u(k+1) \mid b(k) \wedge b(k+\mu+1) = 1$), donc $u(k+1) \mid p(k)$. De même, $u(k-\mu) \mid b(k) \mid q(k)a(k)b(k+1)$, donc $u(k-\mu) \mid q(k)$. Mais alors $u(k+1)$ divise aussi $q(k+\mu+1)$, or $p(k)$ et $q(k+h)$ sont premiers entre eux pour tous $h \in \mathbb{N}^*$: contradiction. \square

2.1.2 Degré des solutions

Au lieu des solutions hypergéométriques de $\Delta T = t$, on s'est ramené à chercher les solutions polynomiales d'une récurrence linéaire à coefficients eux-même polynomiaux. Pour les trouver, il suffit de remplacer l'inconnue par un polynôme générique et d'identifier les coefficients de k^i de part et d'autre... pourvu bien entendu d'avoir une borne sur le degré des solutions, ce à quoi nous allons nous employer.

Théorème 3. Posons

- si $\deg p = \deg q$ et $c = 1$,

$$d = \max(\{\deg r - \deg p + 1, (p(k) - q(k))[k^{\deg p - 1}]\} \cap \mathbb{N});$$

- sinon,

$$d = \deg r - \max(\deg p, \deg q).$$

Le degré des solutions polynomiales de (2) est borné par d .

Démonstration. On distingue deux cas. Si $\deg p(k) \neq \deg q(k)$ ou $c \neq 1$, les termes de plus haut degré dans le membre gauche de (2) ne s'annulent pas. Par conséquent, le degré du membre de gauche est $d + \max(\deg p(k), \deg q(k))$. On a donc :

$$d + \max(\deg p(k), \deg q(k)) = \deg r(k),$$

soit :

$$d = \deg r(k) - \max(\deg p(k), \deg q(k)),$$

ce qui permet de conclure.

Si $\deg p(k) = \deg q(k)$ et $c = 1$, les termes de plus haut degré du membre de gauche de (2) s'annulent. On regarde donc les termes de second plus haut degré :

– s'ils ne s'annulent pas, on obtient par le même calcul que précédemment :

$$d = \deg r(k) - \deg p(k) + 1;$$

– s'ils s'annulent, on pose

$$\begin{aligned} p(k) &= k^n + \beta_p k^{n-1} + \mathcal{O}(k^{n-2}) \\ q(k) &= k^n + \beta_q k^{n-1} + \mathcal{O}(k^{n-2}) \\ s(k) &= \alpha_s k^d + \beta_s k^{d-1} + \mathcal{O}(k^{d-2}). \end{aligned}$$

Alors :

$$\begin{aligned} s(k+1) &= \alpha_s k^d + (\alpha_s d + \beta_s) k^{d-1} + \mathcal{O}(k^{d-2}) \\ p(k)s(k+1) &= \alpha_s k^{d+n} + ((\alpha_s d + \beta_s) + \beta_p \alpha_s) k^{d+n-1} + \mathcal{O}(k^{d-2}) \\ q(k)s(k) &= \alpha_s k^{d+n} + (\alpha_s \beta_q + \beta_s) k^{d+n-1} + \mathcal{O}(k^{n+d-2}). \end{aligned}$$

On obtient donc :

$$cp(k)s(k+1) - q(k)s(k) = \alpha_s (d + \beta_p - \beta_q) k^{d+n-1} + \mathcal{O}(k^{d+n-2}).$$

Par hypothèse, les termes en k^{n+d-1} s'annulent, d'où $d = \beta_q - \beta_p$. \square

2.1.3 Solutions en forme close

L'algorithme de Gosper permet donc de décider si un terme hypergéométrique $t(k)$ admet une primitive discrète hypergéométrique. Nous avons prétendu qu'il nous permettrait de garantir, en cas d'échec, que $\sum_{k=0}^{n-1} t(k)$ n'a pas de formule close. C'est en effet le cas.

Proposition. Supposons que $T(n) = \sum_{k=0}^{n-1} t_k$ admet une formule close. Alors $T(n)$ est la somme d'un terme hypergéométrique et d'une constante.

Démonstration. Posons $T(n) = s_1(n) + \dots + s_m(n) + c$, où c est une constante et les s_i sont des termes hypergéométriques non constants deux à deux non associés. Pour tout i , $s_i(n+1)$, $s_i(n)$ et $d_i(n) = s_i(n+1) - s_i(n)$ sont des termes associés, donc

$$d_1(n) + \dots + d_m(n)$$

est l'unique écriture de $t(n)$ comme somme de termes deux à deux non associés ; donc $m = 1$ et la somme est réduite à $t(n)$ lui-même. \square

2.2 Sommation définie : algorithme de Zeilberger

Considérons la somme

$$f(n) = \sum_k t(n, k)$$

où $t(n, k)$ est hypergéométrique en ses deux arguments.

On cherche à exprimer $f(n)$ sous forme close.

Notons la différence entre le problème abordé ici et celui pour lequel on a étudié l'algorithme de Gosper. L'algorithme de Gosper traite complètement le cas d'une somme indéfinie ($\sum_k t(n, k)$). Si une forme close existe, il la retourne, sinon, il échoue. Celui de Zeilberger traite les sommations définies ($\sum_{k \in \mathbb{Z}} t(n, k)$). Il est clair que pour $t(n, k)$ donné, on peut avoir une solution au second problème sans en avoir une pour le premier (l'analogie avec les calculs d'intégrale est immédiate).

2.2.1 Termes hypergéométriques propres

Définition. Une fonction $t(n, k)$ est un terme hypergéométrique propre s'il existe $p \in \mathbb{K}[n, k]$, $m^\pm \in \mathbb{N}$, $u_\alpha^\pm, v_\alpha^\pm, w_\alpha^\pm \in \mathbb{Z}$, $x \in \mathbb{K}$ tels que

$$t(n, k) = p(n, k) \frac{(u_1^+ n + v_1^+ k + w_1^+)! \dots (u_{m^+}^+ n + v_{m^+}^+ k + w_{m^+}^+)!}{(u_1^- n + v_1^- k + w_1^-)! \dots (u_{m^-}^- n + v_{m^-}^- k + w_{m^-}^-)!} x^k.$$

Insistons sur le fait que les coefficients $u_\alpha^\pm, v_\alpha^\pm$ sont des entiers et non des éléments quelconques de \mathbb{K} . Ils ne peuvent notamment pas contenir de paramètres formels libres⁵.

Tout terme hypergéométrique propre vérifie une relation de récurrence non triviale à coefficients dépendant seulement de n , dont on sait même borner explicitement l'ordre. Plus précisément, soit $t(n, k)$ un terme hypergéométrique propre. Avec les notations de la définition, soient

$$J = \sum_{\alpha, \pm} |v_\alpha^\pm| \quad \text{et} \quad I = \deg p + J \left(\sum_{\alpha, \pm} |u_\alpha^\pm| - 1 \right) + 1.$$

On a alors le résultat suivant.

Lemme. Il existe une famille $(a_{i,j})_{i \in \llbracket 0, I \rrbracket, j \in \llbracket 0, J \rrbracket}$ non nulle de polynômes de $\mathbb{K}[n]$ telle que, pour tous n et k vérifiant

- (i) $p(n, k) \neq 0$ et $\forall \alpha, u_\alpha^- n + v_\alpha^- k + w_\alpha^- \notin \mathbb{Z}_-^*$, i.e. $t(n, k)$ est non nul⁶ ;
- (ii) $\forall \alpha, \forall i \in \llbracket 0, I \rrbracket, \forall j \in \llbracket 0, J \rrbracket, u_\alpha^+ (n - j) + v_\alpha^+ (k - i) + w_\alpha^+ \notin \mathbb{Z}_-^*$, de sorte que tous les numérateurs de (3) sont définis ;

l'on ait

$$\sum_{i=0}^I \sum_{j=0}^J a_{i,j}(n) t(n - j, k - i) = 0. \quad (3)$$

Démonstration. On remarque que si $t(n, k)$ est un polynôme, on obtient en identifiant les coefficients des deux membres de (3), vus comme des polynômes en k à coefficients dans $\mathbb{K}[n]$, un système d'équations linéaires dont les solutions sont précisément les récurrences vérifiées par t . Or on peut aisément se ramener à ce cas en multipliant l'équation par le produit des dénominateurs y apparaissant. Mais on souhaite s'assurer que le système obtenu possédera toujours une solution non triviale, ce qui demande essentiellement un contrôle sur le nombre d'équations, c'est-à-dire sur le degré en k du polynôme.

Soient n et k vérifiant les conditions (i) et (ii). Comme $t(n, k)$ est non nul, (3) équivaut à

$$\sum_{i=0}^I \sum_{j=0}^J a_{i,j}(n) \frac{t(n - j, k - i)}{t(n, k)} = 0.$$

⁵En revanche, pour peu d'arriver à donner un sens à la fonction factorielle sur \mathbb{K} , on pourrait prendre $w_\alpha^\pm \in \mathbb{K}$. Notamment, si \mathbb{K} est un sous-corps de \mathbb{C} , les résultats de cette section tiennent pour $z! = \Gamma(z + 1)$.

⁶On convient que $1/x! = 0$ pour x entier strictement négatif.

Fixons $i \in \llbracket 0, I \rrbracket$ et $j \in \llbracket 0, J \rrbracket$, et considérons un facteur $s(n, k) = (un + vk + w)!$ de $t(n, k)$: on a, puisque u et v sont des entiers

$$\begin{aligned} \frac{s(n-j, k-i)}{s(n, j)} &= \frac{(un + vk + w - uj - ui)!}{(un + vk + w)!} \\ &= \begin{cases} (un + vk + w)^{|uj+vi|} & \text{si } uj + vi < 0 \\ 1/(un + vk + w)^{|uj+vi|} & \text{si } uj + vi > 0. \end{cases} \end{aligned}$$

Ainsi

$$\begin{aligned} \frac{t(n-j, k-i)}{t(n, k)} &= \frac{p(n-j, k-i)}{p(n, k)} \\ &\quad \prod_{\substack{\alpha \in \llbracket 1, m^+ \rrbracket \\ u_\alpha^+ j + v_\alpha^+ i < 0}} (u_\alpha^+ n + v_\alpha^+ k + w_\alpha^+)^{|u_\alpha^+ j + v_\alpha^+ i|} \\ &\quad \times \frac{\prod_{\substack{\alpha \in \llbracket 1, m^+ \rrbracket \\ u_\alpha^+ j + v_\alpha^+ i > 0}} (u_\alpha^+ n + v_\alpha^+ k + w_\alpha^+)^{u_\alpha^+ j + v_\alpha^+ i}}{\prod_{\substack{\alpha \in \llbracket 1, m^- \rrbracket \\ u_\alpha^- j + v_\alpha^- i > 0}} (u_\alpha^- n + v_\alpha^- k + w_\alpha^-)^{u_\alpha^- j + v_\alpha^- i}} \\ &\quad \times \frac{\prod_{\substack{\alpha \in \llbracket 1, m^- \rrbracket \\ u_\alpha^- j + v_\alpha^- i < 0}} (u_\alpha^- n + v_\alpha^- k + w_\alpha^-)^{|u_\alpha^- j + v_\alpha^- i|}}{\prod_{\substack{\alpha \in \llbracket 1, m^- \rrbracket \\ u_\alpha^- j + v_\alpha^- i < 0}} (u_\alpha^- n + v_\alpha^- k + w_\alpha^-)^{|u_\alpha^- j + v_\alpha^- i|}}. \quad (4) \end{aligned}$$

Il n'y a plus qu'à sommer ces expressions suivant i et j , et les réduire au même dénominateur. Dans le premier des deux produits apparaissant au dénominateur, examinons le facteur d'indice α fixé quand i et j varient. Pour tous $i \in \llbracket 0, I \rrbracket$ et $j \in \llbracket 0, J \rrbracket$ vérifiant $uj + vi > 0$, on observe que $(un + vk + w)^{uj+vi}$ divise (comme polynôme en n et k)

$$(un + vk + w)^{\max(uJ, 0) + \max(vI, 0)}.$$

De même, les facteurs de la forme $(un + vk + w)^{|uj+vi|}$ avec $uj + vi > 0$ qui composent le second produit divisent

$$(un + vk + w)^{\max(-uJ, 0) + \max(-vI, 0)}.$$

Un dénominateur commun convenable est donc

$$\begin{aligned} \delta(n, k) = p(n, k) &\quad \prod_{\substack{\alpha \in \llbracket 1, m^+ \rrbracket \\ u_\alpha^+ j + v_\alpha^+ i > 0}} (u_\alpha^+ n + v_\alpha^+ k + w_\alpha^+)^{\max(u_\alpha^+ J, 0) + \max(v_\alpha^+ I, 0)} \\ &\quad \prod_{\substack{\alpha \in \llbracket 1, m^- \rrbracket \\ u_\alpha^- j + v_\alpha^- i < 0}} (u_\alpha^- n + v_\alpha^- k + w_\alpha^-)^{\max(-u_\alpha^- J, 0) + \max(-v_\alpha^- I, 0)}. \quad (5) \end{aligned}$$

Au numérateur de (4) et dans (5), chaque factorielle montante de la forme $(un + vk + w)^{|uj+vi|}$ a, au vu du domaine où varient i et j , un degré en k inférieur

à $|u|J + |v|I$. Il en va de même des factorielles descendantes. En additionnant ces contributions, on voit que $\delta(n, k) \sum_{i=0}^I \sum_{j=0}^J a_{i,j}(n) \frac{t(n-j, k-i)}{t(n, k)}$ est un polynôme en k de degré inférieur à

$$\deg p + I \sum_{\alpha, \pm} |v_{\alpha}^{\pm}| + J \sum_{\alpha, \pm} |u_{\alpha}^{\pm}| = \deg p + IJ + (I + J - \deg p - 1)$$

Or, en identifiant les coefficients des deux membres, on obtient un système linéaire homogène à $(I + 1)(J + 1)$ inconnues $a_{i,j}$. Il y a donc plus d'inconnues que d'équations, et le système admet une solution non triviale. \square

En principe, le lemme précédent donne déjà un algorithme pour trouver un opérateur de récurrence qui annule $\sum_k t(n, k)$. On peut en effet réduire au même dénominateur le membre gauche de (3), puis identifier les coefficients des monômes en k de part et d'autre. Mais cela demande une quantité de calculs qui rend la méthode inutilisable dans la plupart des cas.

Zeilberger est parvenu à améliorer considérablement ce résultat, en montrant que tout terme hypergéométrique propre $t(n, k)$ vérifie une récurrence qui ne fait pas intervenir la variable k . Son idée est de chercher, au lieu d'une récurrence homogène, une relation de récurrence dont les termes du second membre se télescopent quand on en fait la somme selon k .

Théorème 4 (Zeilberger). Sous les mêmes hypothèses que dans le lemme, t est solution d'une récurrence non triviale de la forme

$$\sum_{j=0}^J a_j(n) t(n + j, k) = g(n, k + 1) - g(n, k) \quad (6)$$

où g est un terme hypergéométrique associé à t et $(a_j) \in \mathbb{K}[n]^J$.

Démonstration. D'après le théorème précédent, t est solution d'une récurrence linéaire à coefficients dans $\mathbb{K}[n]$. Il existe donc un polynôme $P \neq 0$ à coefficients dans \mathbb{K} tel que l'opérateur $P(n, N, K)$ annule $t(n, k)$. Choisissons un tel P de degré en K minimal. Effectuons la division « euclidienne » de P par $K - 1$ (qui est unitaire) :

$$P(n, N, K) = P(n, N, 1) - (K - 1)Q(n, N, K).$$

Posant $g(n, k) = Q(n, N, K)t(n, k)$ (on a alors $g(n, k)/t(n, k) \in \mathbb{K}(n, k)$), il vient

$$P(n, N, 1)t(n, k) = (K - 1)Q(n, N, K)t(n, k) = g(n, k + 1) - g(n, k)$$

qui est de la forme cherchée.

Reste à s'assurer que $P(n, N, 1)$ est non nul. Si P ne dépend pas de K , *i.e.* si $P(n, N, 1) = P(n, N, K)$, c'est déjà le cas. Sinon, supposons qu'il le soit. Alors g ne dépend que de n , donc il existe des polynômes $\alpha(n)$ et $\beta(n)$ tels que $g(n + 1)/g(n) = \alpha(n)/\beta(n)$. Ainsi $(\alpha(n)N + \beta(n))Q(n, N, K)$ est un opérateur annihilant $t(n, k)$, ce qui contredit la minimalité de $\deg_K P$. \square

2.2.2 Algorithme de Zeilberger

Le théorème de la section précédente assure l'existence d'une récurrence à coefficients polynomiaux vérifiée par les termes hypergéométriques propres. Dans le but de trouver finalement une formule close pour la somme, on cherche une récurrence d'ordre minimal.

Comment la calculer ? L'idée est de fixer son ordre J et de l'exprimer sous une forme exploitable par l'algorithme de Gosper. On itère le procédé en incrémentant l'ordre de la récurrence tant que l'algorithme échoue. La terminaison est assurée par l'existence d'une borne calculable pour J .

On part donc de :

$$f(n) = \sum_k t(n, k)$$

où $t(n, k)$ est hypergéométrique en ses deux arguments.

Le théorème précédent assure l'existence de la récurrence suivante :

$$\sum_{j=0}^J a_j(n) t(n+j, k) = g(n, k+1) - g(n, k)$$

Posons $T(k) = \sum_{j=0}^J a_j(n) t(n+j, k)$.

On a :

$$\frac{T(k+1)}{T(k)} = \frac{\sum_{j=0}^J a_j(n) \frac{t(n+j, k+1)}{t(n, k+1)}}{\sum_{j=0}^J a_j(n) \frac{t(n+j, k)}{t(n, k)}} \frac{t(n, k+1)}{t(n, k)}$$

L'expression $t(n, k+1)/t(n, k)$ étant rationnelle, on peut la réécrire sous la forme $r_1(n, k)/r_2(n, k)$ avec r_1 et r_2 des polyômes en n et k . On pose également $t(n, k)/t(n-1, k) = s_1(n, k)/s_2(n, k)$ avec s_1 et s_2 également polynômes en n et k .

Par ailleurs :

$$\frac{t(n+j, k)}{t(n, k)} = \prod_{i=0}^{j-1} \frac{t(n+j-i, k)}{t(n+j-i-1, k)} = \prod_{i=0}^{j-1} \frac{s_1(n+j-i, k)}{s_2(n+j-i, k)}$$

$T(k+1)/T(k)$ étant rationnelle, on peut l'écrire sous forme canonique :

$$\begin{aligned} \frac{T(k+1)}{T(k)} &= \frac{\sum_{j=0}^J a_j(n) \prod_{i=0}^{j-1} \frac{s_1(n+j-i, k+1)}{s_2(n+j-i, k+1)} r_1(n, k)}{\sum_{j=0}^J a_j(n) \prod_{i=0}^{j-1} \frac{s_1(n+j-i, k)}{s_2(n+j-i, k)} r_2(n, k)} \\ &= \frac{\sum_{j=0}^J a_j(n) \{ \prod_{i=0}^{j-1} s_1(n+i-j, k+1) \prod_{r=j+1}^J s_2(n+r, k+1) \}}{\sum_{j=0}^J a_j(n) \{ \prod_{i=0}^{j-1} s_1(n+i-j, k) \prod_{r=j+1}^J s_2(n+r, k) \}} \\ &\quad \times \frac{\prod_{r=1}^J s_2(n+r, k)}{\prod_{r=1}^J s_2(n+r, k+1)} \\ &= \frac{p_0(k+1)}{p_0(k)} \frac{u(k)}{v(k)} \end{aligned}$$

avec :

$$p_0(k) = \sum_{j=0}^J a_j(n) \prod_{i=0}^{j-1} s_1(n+i-j, k) \prod_{r=j+1}^J s_2(n+r, k)$$

$$u(k) = r_1(n, k) \prod_{r=1}^J s_2(n+r, k)$$

$$v(k) = r_2(n, k-1) \prod_{r=1}^J s_2(n+r, k)$$

On réécrit l'égalité précédente pour faire apparaître la forme voulue :

$$\frac{u(k)}{v(k)} = c \frac{r_0(k+1)}{r_0(k)} \frac{p(k)}{q(k+1)}$$

avec $\forall h \in \mathbb{N}^*, p(k) \wedge q(k+h) = 1$. On obtient donc :

$$\frac{T(k+1)}{T(k)} = c \frac{r(k+1)}{r(k)} \frac{p(k)}{q(k+1)}$$

avec $\forall h \in \mathbb{N}^*, p(k) \wedge q(k+h) = 1$.

Nous pouvons donc appliquer l'algorithme de Gosper à $T(k)$, qui est sommable si et seulement si :

$$cp(k)s(k+1) - q(k)s(k) = r(k) \quad (7)$$

admet une solution polynomiale $s(k)$.

Posons $s(k) = \sum_{m=0}^d \alpha_m k^m$ où d est déterminé de la manière usuelle par l'algorithme de Gosper.

Remarquons que le membre gauche de cette égalité ne dépend pas des $a_j(n)$ et que le membre droit en est une combinaison linéaire. En remplaçant $s(k)$ par son expression polynomiale dans l'équation (7), on obtient donc un système linéaire en les inconnues $(a_j)_{j \in \llbracket 0..J \rrbracket}$ et $(\alpha_m)_{m \in \llbracket 0..d \rrbracket}$.

Il suffit alors de résoudre ce système. S'il n'admet aucune solution, on incrémente l'ordre J de la récurrence et on reprend l'algorithme. S'il admet une solution, on a obtenu les $(a_j)_{j \in \llbracket 0..J \rrbracket}$, et l'algorithme de Gosper nous indique que :

$$g(n, k) = \frac{q(k)}{r(k)} s(k) T(k).$$

Sommons alors la récurrence sur k . On obtient :

$$a_0(n) f(n) + a_1(n) f(n+1) + \dots + a_J(n) f(n+J) = g(n, A) - g(n, 0)$$

Où A est le plus petit entier h tel que $f(n, h) = 0$ et $f(n, -h) = 0$ pour tout $n \in \llbracket 0..J \rrbracket$. Il ne reste plus qu'à résoudre cette récurrence pour trouver $f(n)$.

3 Formules closes et preuves

Les résultats obtenus jusqu'ici permettent de trouver des récurrences vérifiées par toutes sortes de sommes. On peut utiliser ces récurrences pour démontrer des

identités, et une récurrence du premier ordre se transforme instantanément en formule close. Mais il est possible de trouver des preuves plus simples et surtout plus systématiques, et des formules closes pour des sommes qui nous résistent encore. C'est l'objet de cette dernière section.

3.1 Preuves WZ

On s'intéresse maintenant uniquement au problème de la preuve d'identités. Wilf s'est aperçu qu'une version simplifiée de l'algorithme de Zeilberger permettait de trouver des certificats, c'est-à-dire des preuves systématiques, de forme fixe, pour de nombreuses identités.

On cherche à prouver :

$$\sum_k f(n, k) = u(n)$$

En divisant éventuellement par $u(n)$, on ramène le problème à

$$\sum_k t(n, k) = c$$

où $c = 0$ ou 1 est une constante.

Théorème 5. Soit $((t(n, k))_{k \in \mathbb{N}}, s(n, k)_{k \in \mathbb{N}})$ un couple de termes hypergéométriques tel que :

$$\begin{aligned} t(n+1, k) - t(n, k) &= s(n, k+1) - s(n, k) \\ \forall n \in \mathbb{N} \quad \lim_{k \rightarrow \pm\infty} s(n, k) &= 0 \end{aligned}$$

Alors $\sum_k t(n, k)$ est constante.

Démonstration. Soient $A, B \in \mathbb{N}$ On a :

$$\begin{aligned} \Delta \sum_{k=-B}^A t(n, k) &= \sum_{k=-B}^A (s(n, k+1) - s(n, k)) \\ &= s(n, A+1) - s(n, -B) \end{aligned}$$

D'où, en faisant tendre A et B vers l'infini :

$$\Delta \sum_k t(n, k) = 0$$

soit $\sum_k t(n, k) = c$ avec c constante. □

On appelle le couple (t, s) une *paire WZ*, et on pose $r(n, k) = t(n, k)/s(n, k)$, que l'on appelle *certificat WZ de l'égalité* $\sum_k t(n, k) = c$, car la donnée d'une identité $\sum_k t(n, k) = c$ et d'un certificat $r(n, k)$ permet de vérifier si l'identité est correcte. En effet, le théorème précédent assure que l'identité est vérifiée si $t(n+1, k) - t(n, k) = s(n, k+1) - s(n, k)$.

Maintenant que le principe de la certification est posé, reste à savoir comment trouver le certificat.

Posons $T(n, k) = t(n+1, k) - t(n, k)$. Il est clair que $T(k)$ est hypergéométrique en k . On peut donc lui appliquer l'algorithme de Gosper. Si celui-ci retourne un terme $s(n, k)$, alors $s(n, k)/t(n, k)$ est le certificat recherché. Le théorème suivant caractérise les égalités pour lesquelles l'algorithme WZ prouve une identité.

Théorème 6. Soit $t(n, k)$ hypergéométrique en les deux variables n et k , tel que $T(n) = \sum_k t(n, k)$ existe et soit hypergéométrique en n . Alors l'algorithme de Zeilberger produit une récurrence du premier ordre pour cette somme si et seulement si la méthode WZ prouve l'identité $T(n) = \sum_k t(n, k)$.

Les calculs nécessaires à la recherche d'un certificat WZ sont très vite difficiles à suivre. En revanche, avec un peu de soin, on peut vérifier la preuve à la main. Donnons une preuve compliquée — mais standardisée! — d'une formule simple.

Exemple. Considérons l'identité $\sum_k \binom{n}{k} = 2^n$. On pose

$$t(n, k) = \frac{n!}{2^n k! (n-k)!}.$$

En utilisant un logiciel de calcul formel, on trouve le certificat

$$r(n, k) = \frac{-k}{2(n-k+1)}.$$

Conformément à la procédure standard, on calcule

$$\begin{aligned} & r(n, k+1)t(n, k+1) - r(n, k)t(n, k) \\ &= \frac{n!}{2^n k! (n-k)!} \left(\frac{-(k+1)(n-k)}{2(n-k)(k+1)} - \frac{k}{2(n-k+1)} \right) \\ &= \frac{n!}{2^n k! (n-k)!} \frac{2k-n-1}{2(n-k+1)} \end{aligned}$$

et

$$\begin{aligned} t(n+1, k) - t(n, k) &= \frac{n!}{2^n k! (n-k)!} \left(\frac{n+1}{2(n-k+1)} - 1 \right) \\ &= \frac{n!}{2^n k! (n-k)!} \frac{2k-n-1}{2(n-k+1)}. \end{aligned}$$

On a donc

$$t(n+1, k) - t(n, k) = r(n, k+1)t(n, k+1) - r(n, k)t(n, k)$$

d'où, en sommant selon k , l'identité cherchée.

3.2 Résolution des récurrences : algorithme de Petkovšek

Dans les cas favorables, non seulement l'algorithme de Zeilberger trouve une formule close pour la somme $s(n) = \sum_{k=0}^{n-1} t(n, k)$, mais il donne une preuve en une ligne de l'identité. On s'intéresse ici au cas où la récurrence obtenue est d'ordre deux ou plus, et l'on cherche à savoir si $s(n)$ admet une expression en forme close. Notre démarche va consister à calculer, grâce à l'algorithme Hyper de Petkovšek, une base de l'espace des solutions en forme close de la récurrence, puis à voir si une combinaison linéaire de ces solutions de base peut s'identifier à s .

3.2.1 Calcul d'une solution hypergéométrique

Comme d'habitude, on va ramener la recherche de solutions hypergéométriques au problème qui nous intéresse à celle de solutions polynomiales à des problèmes similaires. Le lemme suivant montre que pour trouver les solutions polynomiales, on peut une fois de plus procéder par identification des coefficients.

Lemme. Soit $L = \sum_{i=0}^I a_i(n)N^i$ ($a_0, a_I \neq 0$) un opérateur de récurrence linéaire d'ordre r à coefficients polynomiaux, et soit b un polynôme. Posons

$$\tilde{a}_j(n) = \sum_{i=j}^I \binom{i}{j} a_i(n) \quad \text{et} \quad \beta = \max_j (\deg \tilde{a}_j - j).$$

Alors les solutions de l'équation $Ly = b$ sont de degré borné par

$$d_0(L) = \max \left(\max(\deg b, -1) - \beta, \max \left\{ x \in \mathbb{N} \mid \sum_{\deg \tilde{a}_j = \beta + j} \lambda(\tilde{a}_j) x^j = 0 \right\} \right).$$

Démonstration. Soit $y(n) = \sum_{k=0}^d \alpha_k n^k$ ($\alpha_d \neq 0$) tel que $Ly = b$. On a $L = \sum_{i=0}^I a_i(n) (\Delta + 1)^i = \sum_{j=0}^I \tilde{a}_j(n) \Delta^j$, donc

$$\sum_{j=0}^I \sum_{k=0}^d \tilde{a}_j(n) \alpha_k \Delta^j n^k = b(n).$$

On vérifie par récurrence sur j que $\Delta^j n^k = k^j n^{k-j} + \mathcal{O}(n^{k-j-1})$. On estime le degré du membre gauche :

$$\deg(Ly(n)) \leq \max_j \deg(\tilde{a}_j(n) \Delta^j y(n)) \leq \max_j ((\deg \tilde{a}_j(n)) + (d - j)) = d + \beta.$$

Si $d + \beta \leq 0$, on a $d \leq d_0$. Si $d + \beta > 0$, deux cas se présentent. Ou bien $Ly(n)$ est de degré exactement $d + \beta$, et $d = \deg b(n) - \beta \leq d_0$. Ou bien

$$(Ly(n))[n^{d+\beta}] = \alpha_d \sum_{\deg \tilde{a}_j = \beta + j} \lambda(\tilde{a}_j) d^j = 0 ;$$

alors d est racine du polynôme $\sum_{\deg \tilde{a}_j = \beta + j} \lambda(\tilde{a}_j) X^j$ (qui par définition de β n'est pas le polynôme nul), et à nouveau $d \leq d_0$. \square

Venons-en à l'algorithme de Petkovšek proprement dit, qui permet de trouver une solution hypergéométrique d'une récurrence linéaire homogène à coefficients polynomiaux. Si $p(n)$ et $q(n)$ sont deux polynômes, on note

$$(p \otimes_I q)_i = \prod_{j=0}^{i-1} p(n+j) \prod_{j=i}^{I-1} q(n+j+1).$$

L'algorithme de la figure 2 prend en entrée un opérateur de récurrence à coefficients polynomiaux L , et renvoie dès que c'est possible un terme hypergéométrique appartenant au noyau de L .

Donnée. Un opérateur de récurrence $L = \sum_{i=0}^I a_i(n)N^i$ ($a_0, a_I \neq 0$).

Résultat. Une solution hypergéométrique y de $Ly = 0$, ou 0 s'il n'existe pas.

- ▶ Pour chaque $p(n)$ unitaire divisant $a_0(n)$
 - ▶ Pour chaque $q(n)$ de unitaire divisant $a_I(n - I)$
 - ▶ Soit $m_{p,q} = \max_{i \in \llbracket 0, I \rrbracket} \deg(a_i(p \otimes_I q)_i)$.
 - ▶ Pour chaque c racine de $\sum_{i=0}^I (a_i(p \otimes_I q)_i)[n^{m_{p,q}}] X^i$
 - ▶ Soit $L_{p,q,c} = \sum_{i=0}^I c^i a_i(n) (p \otimes_I q)_i(n) N^i$.
 - ▶ Chercher par identification des coefficients une solution polynomiale $r(n)$ de l'équation $L_{p,q,c} r(n) = 0$ sous la forme $r(n) = \sum_{i=0}^{d_0(L_{p,q,c})} \alpha_i n^i$, où d_0 est donné par le lemme. En cas de succès
 - ▶ Renvoyer $\prod_{j=1}^n \left(c \frac{p(j)}{q(j+1)} \frac{r(j+1)}{r(j)} \right)$.
- ▶ En cas d'échec d'une étape quelconque, renvoyer 0.

FIG. 2 – Algorithme Hyper

Démonstration. Pour tous p et q , le polynôme $\sum_{i=0}^I \sum_{j=0}^r (a_i(p \otimes_D q)_i)[n^m] X^i$ est non nul par maximalité de m_{pq} , donc chacune des boucles est bornée et le calcul se termine. Il s'agit de montrer d'une part que le résultat est toujours solution de $Ly = 0$, et d'autre part que l'algorithme ne renvoie 0 que s'il n'existe pas de solution non nulle.

Soit $y(n)$ un terme hypergéométrique. Écrivons la fraction rationnelle associée sous forme normale :

$$\frac{y(n+1)}{y(n)} = c \frac{p(n)}{q(n+1)} \frac{r(n+1)}{r(n)}.$$

On a alors les équivalences :

$$\begin{aligned} Ly(n) = 0 &\iff \sum_{i=0}^I a_i(n) \left(\prod_{j=0}^{i-1} \frac{y(n+j+1)}{y(n+j)} \right) y(n) = 0 \\ &\iff \sum_{i=0}^I a_i(n) c^i \left(\prod_{j=0}^{i-1} \frac{p(n+j)}{q(n+j+1)} \right) \frac{r(n+i)}{r(n)} = 0 \\ &\iff \sum_{i=0}^I a_i(n) c^i (p \otimes_I q)_i(n) r(n+i) = 0 \end{aligned}$$

ce qui montre déjà qu'un terme non nul renvoyé par l'algorithme est bien solution de l'équation.

Supposons réciproquement $Ly(n) = 0$, et vérifions que y passe tous les tests présents dans la procédure. Le coefficient de $n^{m_{p,q}}$ dans la dernière équation est exactement

$$\sum_{i=0}^I (a_i(p \otimes_I q)_i) [n^{m_{p,q}}] c^i = 0.$$

Par ailleurs, pour $i \neq 0$, le polynôme $p(n)$ divise $(p \otimes_I q)_i(n)$, d'où

$$p(n) \mid a_0(n) (p \otimes_I q)_0(n) r(n),$$

or $p(n) \wedge r(n) = 1$ et $p(r) \wedge q(n+j+1) = 1$ pour $j \in \mathbb{N}$, donc $p(n) \mid a_0(n)$. Par le même raisonnement, $q(n+I) \mid a_I(n)$. Enfin, d'après le lemme, toutes les solutions de l'équation $L_{p,c,c}r(n) = 0$ sont de degré borné par $d_0(L_{p,q,c})$. \square

Au lieu d'interrompre le calcul dès que l'identification des coefficients donne une solution, on peut naturellement le laisser se poursuivre. L'algorithme Hyper renvoie alors toutes les solutions hypergéométriques (à un scalaire près) de la récurrence.

3.2.2 Espace des solutions

Il se trouve que les solutions en forme close de la récurrence sont non pas des sommes de termes hypergéométriques quelconques, mais des combinaisons linéaires de termes hypergéométriques eux-mêmes dans le noyau de L .

Proposition. Notons H l'ensemble des termes hypergéométriques de $\mathbb{K}^{\mathbb{N}}$. Le sous-espace $\ker L \cap \text{Vect } H$ admet une base contenue dans H .

Démonstration. Soit $t \in \ker L \cap \text{Vect } H$. Écrivons t comme somme de termes hypergéométriques deux à deux non associés : $t = t_1 + \dots + t_m$. On a vu que Lt_i est nul ou hypergéométrique associé à t_i , donc $Lt_1 + \dots + Lt_m$ est l'unique écriture de 0 comme somme de termes hypergéométriques deux à deux non associés. Ainsi $Lt_i = 0$ pour tout i . \square

La situation n'est pas encore tout à fait satisfaisante. On n'a même pas de borne sur le nombre de solutions hypergéométriques à additionner pour trouver une formule close qui coïncide avec la somme de départ. Ce qu'il nous manque est un analogue du théorème de Cauchy-Lipschitz⁷, qui donnerait la structure du noyau de L . Mais l'on rencontre un peu le même problème qu'à la section 1.2 : comme les coefficients de L peuvent s'annuler, son noyau peut être de dimension (finie) quelconque.

La solution est similaire : on peut se contenter de chercher des termes vérifiant la récurrence à partir d'un certain rang. Formellement, on considère des « suites » appartenant à l'algèbre quotient $\mathcal{S} = \mathbb{K}^{\mathbb{N}} / \mathbb{K}^{(\mathbb{N})}$. L'opérateur N passe au quotient (il existe un unique endomorphisme \bar{N} de \mathcal{S} tel que $\overline{Nt} = \bar{N}\bar{t}$ pour toute suite t , où la barre désigne la surjection canonique de $\mathbb{K}^{\mathbb{N}}$ dans \mathcal{S}), de sorte que tous les opérateurs de récurrence sur $\mathbb{K}^{\mathbb{N}}$ peuvent être vus naturellement comme des opérateurs de récurrence sur \mathcal{S} .

⁷Pour poursuivre l'analogie avec les équations différentielles, il y a une façon de trouver une base de l'espace de $\ker L \cap \text{Vect } H$ plus efficace que chercher toutes les solutions hypergéométriques : utiliser la « variation de la constante », c'est-à-dire chercher une première solution, s'en servir pour diminuer le degré de la récurrence et trouver une solution indépendante, et ainsi de suite.

Théorème 7. Soit $L = \sum_{j=0}^J a_j(n)N^j$ un opérateur de récurrence linéaire sur \mathcal{S} , d'ordre J , à coefficients polynomiaux. Alors $\dim \ker L = J$.

Le résultat tient toujours si les coefficients sont des suites quelconques, pourvu que a_0 et a_J soient inversibles dans \mathcal{S} .

Démonstration. Soient $y_1, \dots, y_{J+1} \in \mathbb{K}^{\mathbb{N}}$ des représentants d'éléments de $\ker L$. Comme le coefficient a_0 est inversible, il ne s'annule qu'un nombre fini de fois. Pour tout $i \in \llbracket 1, J+1 \rrbracket$, on a $Ly_i(n) = 0$ presque pour tout n . Ainsi, pour n grand, disons $n \geq n_0$,

$$\forall i \in \llbracket 1, J+1 \rrbracket, \quad y_i(n) = - \sum_{j=1}^J \frac{a_j(n)}{a_0(n)} y_i(n+j) = 0 \quad (8)$$

$$\text{et } \begin{bmatrix} y_1(n) \\ \vdots \\ y_{J+1}(n) \end{bmatrix}, \dots, \begin{bmatrix} y_1(n+J) \\ \vdots \\ y_{J+1}(n+J) \end{bmatrix} \text{ sont liés,} \quad (9)$$

de même donc que

$$\begin{bmatrix} y_1(n) \\ \vdots \\ y_{J+1}(n+J) \end{bmatrix}, \dots, \begin{bmatrix} y_{J+1}(n) \\ \vdots \\ y_{J+1}(n+J) \end{bmatrix}.$$

Soit $n \geq n_0$, et écrivons la relation de liaison au rang $n+1$:

$$\forall j \in \llbracket 1, J+1 \rrbracket, \quad \sum_{i=1}^{J+1} \alpha_i y_i(n+j) = 0.$$

On en déduit grâce à (8) :

$$\sum_{i=1}^{J+1} \alpha_i y_i(n) = \sum_{j=1}^J \frac{a_j(n)}{a_0(n)} \sum_{i=1}^{J+1} \alpha_i y_i(n+j) = 0,$$

ce qui montre que la relation valait déjà au rang n . Donc la famille (y_1, \dots, y_{J+1}) est liée, et $\dim \ker L \leq J$.

D'autre part, soit n_1 tel que $a_0(n)$ et $a_J(n)$ ne s'annulent plus pour $n \geq n_1$. Posons

$$\begin{bmatrix} y_0(n_1) & \dots & y_0(n_1+J-1) \\ \vdots & & \vdots \\ y_{J-1}(n_1) & \dots & y_{J-1}(n_1+J-1) \end{bmatrix} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \quad (10)$$

et complétons par $y_i(n)$, $n \geq n_1 + J$ de sorte que les $y_i \in \ker L$ (ce qui est possible puisque $a_J(n) \neq 0$). Par le même raisonnement que ci-dessus, si la famille (y_1, \dots, y_{J-1}) était liée, les lignes de (10) le seraient aussi. Finalement $\dim \ker L = J$. \square

Indiquons enfin, sans entrer dans les détails, que l'algorithme de Petkovšek peut facilement être adapté aux récurrences avec second membre. Cela permet d'appliquer l'algorithme de Zeilberger même quand le terme sous la somme ne s'annule pas en-dehors de l'intervalle de sommation.

Conclusion

Récapitulons les principaux résultats :

- le problème de la sommation indéfinie en forme close « $\sum^n t(k) = ?$ » des termes hypergéométriques propres est algorithmique, c'est-à-dire qu'il peut être résolu par une méthode systématique, sans aucune imagination ;
- il en va de même du problème de la sommation définie en forme close « $\sum_k t(n, k) = ?$ » des termes hypergéométriques propres ;
- et les identités obtenues sont (plus ou moins) aisées à vérifier à la main ; dans les cas favorables, le certificat se résume à une fraction rationnelle et sa vérification à un calcul de routine.

Ces questions ne sont pas seulement décidables : les algorithmes en jeu sont assez simples et assez efficaces pour fonctionner sur des problèmes réels, et ils sont utilisés dans les logiciels de calcul formel majeurs.

Nous sommes aussi équipés pour décider de l'égalité de deux sommes : tout d'abord, si elles s'écrivent en forme close, les propriétés d'unicité des termes hypergéométriques sont telles qu'il suffit pratiquement de comparer leurs formules closes. Si elles n'en ont pas, il est encore possible de chercher si elles vérifient une même relation de récurrence.

En revanche, les preuves que l'on obtient sont des preuves *ad hoc*, considérablement moins éclairantes que les preuves traditionnelles d'identités combinatoires, dont certaines ont donné naissance à des pans entiers de théorie — parmi lesquels, soit dit en passant, la méthode de Zeilberger elle-même. Pourtant, ces preuves ne sont pas aussi vides d'enseignements qu'elles le paraissent au premier abord. Wilf et Zeilberger ont en effet constaté que chaque preuve WZ démontrait, outre l'identité pour laquelle elle avait été calculée, au moins deux autres formules, appelées identité compagnon et identité duale. Ces deux identités « offertes » sont souvent inconnues, parfois intéressantes et rarement triviales. Quand une identité est un cas particulier d'une autre, ni le passage au compagnon ni la dualisation ne commutent en général avec la spécialisation, ce qui peut révéler encore d'autres identités nouvelles. Un pas vers la recherche automatique de théorèmes ?

Références

- [1] Marko PETKOVŠEK, Herbert S. WILF, Doron ZEILBERGER. *A = B*. A K Peters, Ltd., Wellesley, MA, 1996.
- [2] Ronald R. GRAHAM, Donald E. KNUTH, Oren PATASHNIK. *Mathématiques concrètes*. 2^e édition. ITS France, Paris, 1998.