Précision p-adique, exemple et applications

séminaire des doctorants en théorie des nombres de Bordeaux

Tristan Vaccon

Université de Rennes I

15 novembre 2013

Besoin de calculs p-adiques

Pourquoi travailler en p-adique ?

- Passer de $\mathbb{F}_p$ à $\mathbb{Z}_p$ pour revenir à $\mathbb{F}_p$ permet plus de calculs ;

Besoin de calculs p-adiques

Pourquoi travailler en p-adique ?

- Passer de $\mathbb{F}_p$ à $\mathbb{Z}_p$ pour revenir à $\mathbb{F}_p$ permet plus de calculs ;
- Travailler dans $\mathbb{Q}_p$ plutôt que dans $\mathbb{Q}$ peut permettre de gérer efficacement les problèmes d'explosion de taille des coefficients ;

Besoin de calculs p-adiques

Pourquoi travailler en p-adique ?

- Passer de $\mathbb{F}_p$ à $\mathbb{Z}_p$ pour revenir à $\mathbb{F}_p$ permet plus de calculs ;
- Travailler dans $\mathbb{Q}_p$ plutôt que dans $\mathbb{Q}$ peut permettre de gérer efficacement les problèmes d'explosion de taille des coefficients ;
- Certains algorithmes sont p-adiques par nature.

Un exemple d'algorithme p-adique

Factorisation par Hensel

On cherche à factoriser $Q \in \mathbb{Z}[X]$:

Un exemple d'algorithme p-adique

Factorisation par Hensel

On cherche à factoriser $Q \in \mathbb{Z}[X]$:

1. Choisir un p "adapté" au problème ;

Un exemple d'algorithme p-adique

Factorisation par Hensel

On cherche à factoriser $Q \in \mathbb{Z}[X]$:

1. Choisir un p "adapté" au problème ;
2. Factoriser $\overline{Q} \in \mathbb{Z}/p\mathbb{Z}[X]$;

Un exemple d'algorithme p-adique

Factorisation par Hensel

On cherche à factoriser $Q \in \mathbb{Z}[X]$:

1. Choisir un p "adapté" au problème ;
2. Factoriser $\overline{Q} \in \mathbb{Z}/p\mathbb{Z}[X]$;
3. Remonter les facteurs à $\mathbb{Z}/p^k\mathbb{Z}[X]$ (par *remontée de Hensel*) ;

Un exemple d'algorithme p-adique

Factorisation par Hensel

On cherche à factoriser $Q \in \mathbb{Z}[X]$:

1. Choisir un p "adapté" au problème ;
2. Factoriser $\overline{Q} \in \mathbb{Z}/p\mathbb{Z}[X]$;
3. Remonter les facteurs à $\mathbb{Z}/p^k\mathbb{Z}[X]$ (par *remontée de Hensel*) ;
4. Si p^k est assez grand (*borne de Mignotte*), on obtient une décomposition de Q (modulo un travail de recombinaison des facteurs).

Un autre exemple d'algorithme p-adique

Idées de l'algorithme de Kedlaya

Soit C une courbe hyperelliptique de genre g sur $\mathbb{F}_p$, donnée par $y^2 = P(x)$ (avec P de degré $2g+1$, sans facteur carré). On veut connaître $|Jac(C, \mathbb{F}_p)|$.

Un autre exemple d'algorithme p-adique

Idées de l'algorithme de Kedlaya

Soit C une courbe hyperelliptique de genre g sur $\mathbb{F}_p$, donnée par $y^2 = P(x)$ (avec P de degré $2g+1$, sans facteur carré). On veut connaître $|Jac(C, \mathbb{F}_p)|$.

- On note F le Frobenius de $\mathbb{F}_p$. Alors F agit comme un endomorphisme de $H^1_{MW}(C, A)$, cohomologie de Monsky-Washnitzer à coefficient dans A.

Un autre exemple d'algorithme p-adique

Idées de l'algorithme de Kedlaya

Soit C une courbe hyperelliptique de genre g sur $\mathbb{F}_p$, donnée par $y^2 = P(x)$ (avec P de degré $2g+1$, sans facteur carré). On veut connaître $|Jac(C, \mathbb{F}_p)|$.

- On note F le Frobenius de $\mathbb{F}_p$. Alors F agit comme un endomorphisme de $H^1_{MW}(C, A)$, cohomologie de Monsky-Washnitzer à coefficient dans A.
- On prend $A = \mathbb{Z}_p^\dagger[[x, y]]/(P)$. Alors $|Jac(C, \mathbb{F}_p)| = \chi_F(1)$.

Un autre exemple d'algorithme p-adique

Idées de l'algorithme de Kedlaya

Soit C une courbe hyperelliptique de genre g sur $\mathbb{F}_p$, donnée par $y^2 = P(x)$ (avec P de degré $2g+1$, sans facteur carré). On veut connaître $|Jac(C, \mathbb{F}_p)|$.

- On note F le Frobenius de $\mathbb{F}_p$. Alors F agit comme un endomorphisme de $H^1_{MW}(C, A)$, cohomologie de Monsky-Washnitzer à coefficient dans A.
- On prend $A = \mathbb{Z}_p^\dagger[[x, y]]/(P)$. Alors $|Jac(C, \mathbb{F}_p)| = \chi_F(1)$.
- On veut calculer l'action de F sur A :

Un autre exemple d'algorithme p-adique

Idées de l'algorithme de Kedlaya

Soit C une courbe hyperelliptique de genre g sur $\mathbb{F}_p$, donnée par $y^2 = P(x)$ (avec P de degré $2g+1$, sans facteur carré). On veut connaître $|Jac(C, \mathbb{F}_p)|$.

- On note F le Frobenius de $\mathbb{F}_p$. Alors F agit comme un endomorphisme de $H^1_{MW}(C, A)$, cohomologie de Monsky-Washnitzer à coefficient dans A.
- On prend $A = \mathbb{Z}_p^\dagger[[x, y]]/(P)$. Alors $|Jac(C, \mathbb{F}_p)| = \chi_F(1)$.
- On veut calculer l'action de F sur A :

$$F(x) = x^p \mod p \qquad F(y) = y^p \mod p$$
$$P(F(x)) = F(y)^2$$

- Par les conjectures de Weil, $\chi_F \in \mathbb{Z}[T]$, et $|a_i| \leqslant 2^{2g}\sqrt{q}^i$.

1. Précision p-adique et algorithme de Gauss
   - Puissance de la précision p-adique
   - Un exemple : analyse de la perte de précision dans l'échelonnement de Gauss
2. Polygone de Newton
   - Première présentation
   - Factorisation par les pentes
3. Calcul différentiel
   - Le lemme de Caruso
   - Equations différentielles linéaires, et applications

UNIVERSITÉ DE RENNES 1

Plan de l'exposé

Définition de la précision

p-adiques à précision fixée

Les éléments de $\mathbb{Q}_p$ sont de la forme $\sum_{i=-l}^{+\infty} a_i p^i$, avec $a_i \in [\![0, p-1]\!]$, $l \in \mathbb{Z}$ et p un nombre premier.
Nécessairement, en machine, on se restreint au début de ce développement en série, et on ne considère que des éléments de la forme suivante : $\sum_{i=l}^{d-1} a_i p^i + O(p^d)$, avec $l \in \mathbb{Z}$.

Définition de la précision

p-adiques à précision fixée

Les éléments de $\mathbb{Q}_p$ sont de la forme $\sum_{i=-l}^{+\infty} a_i p^i$, avec $a_i \in [\![0, p-1]\!]$, $l \in \mathbb{Z}$ et p un nombre premier.
Nécessairement, en machine, on se restreint au début de ce développement en série, et on ne considère que des éléments de la forme suivante : $\sum_{i=l}^{d-1} a_i p^i + O(p^d)$, avec $l \in \mathbb{Z}$.

Définition

L'**ordre**, ou la **précision absolue** de $\sum_{i=k}^{d-1} a_i p^i + O(p^d)$ est d. Sa **précision relative** correspond au nombre de ses chiffres significatifs, et est donnée par $d - \min\{i \in \mathbb{Z}, a_i \neq 0\}$.

Définition de la précision

p-adiques à précision fixée

Les éléments de $\mathbb{Q}_p$ sont de la forme $\sum_{i=-l}^{+\infty} a_i p^i$, avec $a_i \in [\![0, p-1]\!]$, $l \in \mathbb{Z}$ et p un nombre premier.
Nécessairement, en machine, on se restreint au début de ce développement en série, et on ne considère que des éléments de la forme suivante : $\sum_{i=l}^{d-1} a_i p^i + O(p^d)$, avec $l \in \mathbb{Z}$.

Définition

L'**ordre**, ou la **précision absolue** de $\sum_{i=k}^{d-1} a_i p^i + O(p^d)$ est d. Sa **précision relative** correspond au nombre de ses chiffres significatifs, et est donnée par $d - \min\{i \in \mathbb{Z}, a_i \neq 0\}$.

Exemples

L'ordre de $3 * 7^{-1} + 4 * 7^0 + 5 * 7^1 + 6 * 7^2 + O(7^3)$ est 3, et sa précision relative $4 = 3 - (-1)$.

Précision p-adique contre précision réelle

L'essence de l'étude de la précision p-adique provient de la remarque suivante :

Proposition (Les erreurs de précision ne se cumulent pas)

On a :

$$(a + O(p^k)) + (b + O(p^k)) = a + b + O(p^k).$$

Autrement dit, si a et b sont connus avec une précision $O(p^k)$, alors $a + b$ aussi.

Précision p-adique contre précision réelle

L'essence de l'étude de la précision p-adique provient de la remarque suivante :

Proposition (Les erreurs de précision ne se cumulent pas)

On a :

$$(a + O(p^k)) + (b + O(p^k)) = a + b + O(p^k).$$

Autrement dit, si a et b sont connus avec une précision $O(p^k)$, alors $a + b$ aussi.

Précision p-adique contre précision réelle

L'essence de l'étude de la précision p-adique provient de la remarque suivante :

Proposition (Les erreurs de précision ne se cumulent pas)

On a :

$$(a + O(p^k)) + (b + O(p^k)) = a + b + O(p^k).$$

Autrement dit, si a et b sont connus avec une précision $O(p^k)$, alors $a + b$ aussi.

Remarque

Ceci est le contraire du cas réel, et de ses **erreurs d'arrondis** :

$$(1 + 5 * 10^{-2}) + (2 + 6 * 10^{-2}) = 3 + 1 * 10^{-1} + 1 * 10^{-2}.$$

Autrement dit, si a et b sont connus avec une précision 10^{-n}, alors $a + b$ est connu à $10^{(-n+1)}$.

Précision p-adique contre précision réelle

L'essence de l'étude de la précision p-adique provient de la remarque suivante :

Proposition (Les erreurs de précision ne se cumulent pas)

On a :

$$(a + O(p^k)) + (b + O(p^k)) = a + b + O(p^k).$$

Autrement dit, si a et b sont connus avec une précision $O(p^k)$, alors $a + b$ aussi.

Remarque

Ceci est le contraire du cas réel, et de ses **erreurs d'arrondis** :

$$(1 + 5 * 10^{-2}) + (2 + 6 * 10^{-2}) = 3 + 1 * 10^{-1} + 1 * 10^{-2}.$$

Autrement dit, si a et b sont connus avec une précision 10^{-n}, alors $a + b$ est connu à $10^{(-n+1)}$.

Formules de précision

Proposition (addition)

$$(x_0 + O(p^{k_0})) + (x_1 + O(p^{k_1})) = x_0 + x_1 + O(p^{min(k_0,k_1)})$$

Formules de précision

Proposition (addition)

$$(x_0 + O(p^{k_0})) + (x_1 + O(p^{k_1})) = x_0 + x_1 + O(p^{min(k_0,k_1)})$$

Proposition (multiplication)

$$(x_0 + O(p^{k_0})) * (x_1 + O(p^{k_1})) = x_0 * x_1 + O(p^{min(k_0 - v_p(x_1), k_1 - v_p(x_0))})$$

Formules de précision

Proposition (addition)

$$(x_0 + O(p^{k_0})) + (x_1 + O(p^{k_1})) = x_0 + x_1 + O(p^{min(k_0,k_1)})$$

Proposition (multiplication)

$$(x_0 + O(p^{k_0})) * (x_1 + O(p^{k_1})) = x_0 * x_1 + O(p^{min(k_0 - v_p(x_1), k_1 - v_p(x_0))})$$

Proposition (division)

$$\frac{xp^a + O(p^b)}{yp^c + O(p^d)} = x * y^{-1} p^{a-c} + O(p^{min(d+a-2c, b-c)})$$

En particulier,

$$\frac{1}{yp^c + O(p^d)} = y^{-1} p^{-c} + O(p^{d-2c})$$

Plan de l'exposé

Méthode de Gauss

Échelonnement à la Gauss

Soit

$$M = \begin{bmatrix} m_{1,1} + O(p^k) & m_{1,2} + O(p^k) & \cdots & m_{1,m} + O(p^k) \\ m_{2,1} + O(p^k) & m_{2,2} + O(p^k) & \cdots & m_{2,m} + O(p^k) \end{bmatrix}$$

On suppose

$$\det\left(\begin{bmatrix} m_{1,1} + O(p^k) & m_{1,2} + O(p^k) \\ m_{2,1} + O(p^k) & m_{2,2} + O(p^k) \end{bmatrix}\right) \neq O(p^k).$$

Méthode de Gauss

Échelonnement à la Gauss

$$M \simeq \begin{bmatrix} p^{a_1} + O(p^k) & \cdots & m^{(1)}_{1,m} + O(p^k) \\ O(p^k) & m^{(1)}_{2,2} + O(p^k) & \cdots & m^{(1)}_{2,m} + O(p^k) \end{bmatrix} \qquad \begin{array}{l} L_1 \leftarrow c_1^{-1} L_1 \\ L_2 \leftarrow L_2 - \frac{m_{2,1}}{m_{1,1}} L_1 \end{array}$$

On prend comme pivot le coefficient sur la première colonne de **plus petite valuation**, mis sur la première ligne par un échange de lignes : $M_{1,1} = c_1 * p^{a_1} + O(p^k)$.

De plus, comme $a_1 \leqslant val(m_{j,1})$, $\frac{m_{2,1}}{m_{1,1}}$ est dans $\mathbb{Z}_p$, connu à l'ordre k au moins.

Méthode de Gauss

Échelonnement à la Gauss

$$M \simeq \begin{bmatrix} p^{a_1} + O(p^k) & \cdots & m^{(1)}_{1,m} + O(p^k) \\ \boxed{O(p^k)} & m^{(1)}_{2,2} + O(p^k) & \cdots & m^{(1)}_{2,m} + O(p^k) \end{bmatrix} \qquad \begin{array}{l} L_1 \leftarrow c_1^{-1} L_1 \\ L_2 \leftarrow L_2 - \boxed{\frac{m_{2,1}}{m_{1,1}}} L_1 \end{array}$$

On prend comme pivot le coefficient sur la première colonne de **plus petite valuation**, mis sur la première ligne par un échange de lignes : $M_{1,1} = c_1 * p^{a_1} + O(p^k)$.

De plus, comme $a_1 \leqslant val(m_{j,1})$, $\boxed{\frac{m_{2,1}}{m_{1,1}}}$ est dans $\mathbb{Z}_p$, connu à l'ordre k au moins.

Méthode de Gauss

Échelonnement à la Gauss

On est alors ramenés au cas d'une matrice échelonnée à $O(p^k)$ près :

$$M = \begin{bmatrix} p^{a_1} + O(p^k) & m_{1,2} + O(p^k) & \cdots & m_{1,m} + O(p^k) \\ O(p^k) & p^{a_2} + O(p^k) & \cdots & m_{2,m} + O(p^k) \end{bmatrix}$$

Méthode de Gauss

Échelonnement à la Gauss

On est alors ramenés au cas d'une matrice échelonnée à $O(p^k)$ près :

$$M = \begin{bmatrix} p^{a_1} + O(p^k) & m_{1,2} + O(p^k) & \cdots & m_{1,m} + O(p^k) \\ O(p^k) & p^{a_2} + O(p^k) & \cdots & m_{2,m} + O(p^k) \end{bmatrix}$$

On note que le coefficient $M_{2,2}$ est non nul modulo p^k par hypothèse.

Méthode de Gauss

Échelonnement à la Gauss

$$M \simeq \begin{bmatrix} p^{a_1} + O(p^k) & m_{1,2} + O(p^k) & \cdots & m_{1,m} + O(p^k) \\ \boxed{0} & p^{a_2} + O(p^{k-a_1}) & \cdots & m_{2,m} + O(p^{k-a_1}) \end{bmatrix}$$

$$L_2 \leftarrow L_2 - \frac{M_{2,1}^{(n-1)}}{M_{1,1}^{(n-1)}} L_1$$

En effet, $M_{2,1}^{(n-1)} - \frac{M_{2,1}^{(n-1)}}{M_{1,1}^{(n-1)}} * M_{1,1}^{(n-1)} = 0$ (formellement).

Par ailleurs, $\frac{M_{2,1}^{(n-1)}}{M_{1,1}^{(n-1)}} = \frac{O(p^k)}{p^{a_1} + O(p^k)} = O(p^{k-a_1})$, donc $L_2 - \frac{M_{2,1}^{(n-1)}}{M_{1,1}^{(n-1)}} L_1 = L_2 + O(p^{k-a_1}) L_1$.

Méthode de Gauss

Échelonnement à la Gauss

$$M \simeq \begin{bmatrix} p^{a_1} + O(p^k) & m_{1,2} + O(p^k) & \cdots & m_{1,m} + O(p^k) \\ \boxed{0} & p^{a_2} + O(p^{k-a_1}) & \cdots & m_{2,m} + O(p^{k-a_1}) \end{bmatrix}$$

$$L_2 \leftarrow L_2 + O(p^{k-a_1})L_1$$

En effet, $M_{2,1}^{(n-1)} - \dfrac{M_{2,1}^{(n-1)}}{M_{1,1}^{(n-1)}} * M_{1,1}^{(n-1)} = 0$ (formellement).

Par ailleurs, $\dfrac{M_{2,1}^{(n-1)}}{M_{1,1}^{(n-1)}} = \dfrac{O(p^k)}{p^{a_1}+O(p^k)} = O(p^{k-a_1})$, donc $L_2 - \dfrac{M_{2,1}^{(n-1)}}{M_{1,1}^{(n-1)}} L_1 = L_2 + O(p^{k-a_1})L_1$.

Méthode de Gauss

Échelonnement à la Gauss

À la fin, on obtient :

$$M = \begin{bmatrix} p^{a_1} + O(p^k) & m_{1,2} + O(p^k) & \cdots & m_{1,m} + O(p^k) \\ \boxed{0} & p^{a_2} + O(p^{k-a_1}) & \cdots & \boxed{m_{2,m} + O(p^{k-a_1})} \end{bmatrix}$$

La perte de précision sur la deuxième ligne est $\boxed{a_1}$.

Méthode de Gauss

Échelonnement à la Gauss

À la fin, on obtient :

$$M = \begin{bmatrix} p^{a_1} + O(p^k) & m_{1,2} + O(p^k) & \cdots & m_{1,m} + O(p^k) \\ 0 & p^{a_2} + O(p^{k-a_1}) & \cdots & m_{2,m} + O(p^{k-a_1}) \end{bmatrix}$$

La perte de précision sur la deuxième ligne est a_1 .

Or, $val(\det\left(\begin{bmatrix} m_{1,1} + O(p^k) & m_{1,2} + O(p^k) \\ m_{2,1} + O(p^k) & m_{2,2} + O(p^k) \end{bmatrix}\right)) = a_1 + a_2$, et les a_i sont positifs.

La perte de précision est donc majorée par $val(det((M_{i,j})_{1 \leqslant i \leqslant 2, 1 \leqslant j \leqslant 2}))$

Méthode de Gauss : un résultat

Théorème

Méthode de Gauss : un résultat

Théorème

Soit M une matrice $n \times m$ dont les coefficients sont des entiers p-adiques connus avec la précision uniforme $O(p^k)$, et telles que son mineur principal $\Delta = det((M_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant n})$ vérifie $val(\Delta) < k$.

Méthode de Gauss : un résultat

Théorème

Soit M une matrice $n \times m$ dont les coefficients sont des entiers p-adiques connus avec la précision uniforme $O(p^k)$, et telles que son mineur principal $\Delta = det((M_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant n})$ vérifie $val(\Delta) < k$.
Alors, la perte maximale de précision lorsque l'on effectue l'échelonnement de M par l'algorithme de Gauss peut être majorée par $val(\Delta)$ (autrement dit, on peut calculer les coefficients de la forme échelonnée à précision $O(p^{k-val(\Delta)})$.

Le problème de la détermination de l'indice d'une ligne

Souci des tests à zéro

Un souci majeur apparait lorsqu'on ne considère que des données connues à précision fixée : celui de ne pas savoir si une colonne est sans pivot ou si la précision est insuffisante.

Indétermination de l'indice d'une ligne

$$\begin{bmatrix} 1+O(p^k) & 1+O(p^k) & 1+O(p^k) & 0 \\ 1+O(p^k) & 1+O(p^k) & 0 & 1+O(p^k) \end{bmatrix}$$

$$L_2 \leftarrow L_2 - \frac{M_{2,1}}{M_{1,1}} L_1$$

Le problème de la détermination de l'indice d'une ligne

Souci des tests à zéro

Un souci majeur apparait lorsqu'on ne considère que des données connues à précision fixée : celui de ne pas savoir si une colonne est sans pivot ou si la précision est insuffisante.

Indétermination de l'indice d'une ligne

$$\begin{bmatrix} 1+O(p^k) & 1+O(p^k) & 1+O(p^k) & 0 \\ 0 & O(p^k) & -1+O(p^k) & 1+O(p^k) \end{bmatrix} \qquad L_2 \leftarrow L_2 - (1+O(p^k))L_1$$

Le problème de la détermination de l'indice d'une ligne

Souci des tests à zéro

Un souci majeur apparait lorsqu'on ne considère que des données connues à précision fixée : celui de ne pas savoir si une colonne est sans pivot ou si la précision est insuffisante.

Indétermination de l'indice d'une ligne

$$\begin{bmatrix} 1+O(p^k) & 1+O(p^k) & 1+O(p^k) & 0 \\ 0 & O(p^k) & -1+O(p^k) & 1+O(p^k) \end{bmatrix} \qquad L_2 \leftarrow L_2 - (1+O(p^k))L_1$$

Quel est l'indice de la première ligne ?

Plan de l'exposé

Polygone de Newton d'un polynôme

Définition

Soit $f(X) = a_0 + \cdots + a_n X^n \in K[x]$ un polynôme, avec K un corps discrétement valué. Soit $U = \{(0, v(a_0)), \ldots, (n, v(a_n))\}$.

Polygone de Newton d'un polynôme

Définition

Soit $f(X) = a_0 + \cdots + a_n X^n \in K[x]$ un polynôme, avec K un corps discrétement valué. Soit $U = \{(0, v(a_0)), \ldots, (n, v(a_n))\}$.
On définit *le polygone de Newton* de f comme l'ensemble des points d'abscisse dans $[0, n]$ qui sont au-dessus de l'enveloppe convexe "inférieure" de U, ainsi que cette enveloppe convexe.

Polygone de Newton d'un polynôme

Définition

Soit $f(X) = a_0 + \cdots + a_n X^n \in K[x]$ un polynôme, avec K un corps discrétement valué. Soit $U = \{(0, v(a_0)), \ldots, (n, v(a_n))\}$.
On définit *le polygone de Newton* de f comme l'ensemble des points d'abscisse dans $[0, n]$ qui sont au-dessus de l'enveloppe convexe "inférieure" de U, ainsi que cette enveloppe convexe.
On entend par enveloppe convexe "inférieure" l'enveloppe convexe des points de U, à l'exception de ceux qui sont au-dessus du segment joignant $(0, (a_0))$ à $(n, v(a_n))$.

Polygone de Newton d'un polynôme

Définition

Soit $f(X) = a_0 + \cdots + a_n X^n \in K[x]$ un polynôme, avec K un corps discrétement valué. Soit $U = \{(0, v(a_0)), \ldots, (n, v(a_n))\}$.
On définit *le polygone de Newton* de f comme l'ensemble des points d'abscisse dans $[0, n]$ qui sont au-dessus de l'enveloppe convexe "inférieure" de U, ainsi que cette enveloppe convexe.
On entend par enveloppe convexe "inférieure" l'enveloppe convexe des points de U, à l'exception de ceux qui sont au-dessus du segment joignant $(0, (a_0))$ à $(n, v(a_n))$.
Cela correspond aussi au graphe de la plus grande fonction convexe sur $[0, n]$, notée $Newt_f$, qui soit "sous" les points de U, c'est-à-dire telle que $Newt_f$ est convexe et vérifie pour tout i, $f(i) \leqslant v(a_i)$, et elle est la plus grande (au sens d'être plus grand en chaque point) à le vérifier.

Polygone de Newton d'un polynôme

Définition

Soit $f(X) = a_0 + \cdots + a_n X^n \in K[x]$ un polynôme, avec K un corps discrétement valué. Soit $U = \{(0, v(a_0)), \ldots, (n, v(a_n))\}$.
On définit *le polygone de Newton* de f comme l'ensemble des points d'abscisse dans $[0, n]$ qui sont au-dessus de l'enveloppe convexe "inférieure" de U, ainsi que cette enveloppe convexe.
On entend par enveloppe convexe "inférieure" l'enveloppe convexe des points de U, à l'exception de ceux qui sont au-dessus du segment joignant $(0, (a_0))$ à $(n, v(a_n))$.
Cela correspond aussi au graphe de la plus grande fonction convexe sur $[0, n]$, notée $Newt_f$, qui soit "sous" les points de U, c'est-à-dire telle que $Newt_f$ est convexe et vérifie pour tout i, $f(i) \leqslant v(a_i)$, et elle est la plus grande (au sens d'être plus grand en chaque point) à le vérifier.

Remarque

La même définition fonctionnerait avec $f \in K[[x]]$.

Un exemple

Un polygone de Newton

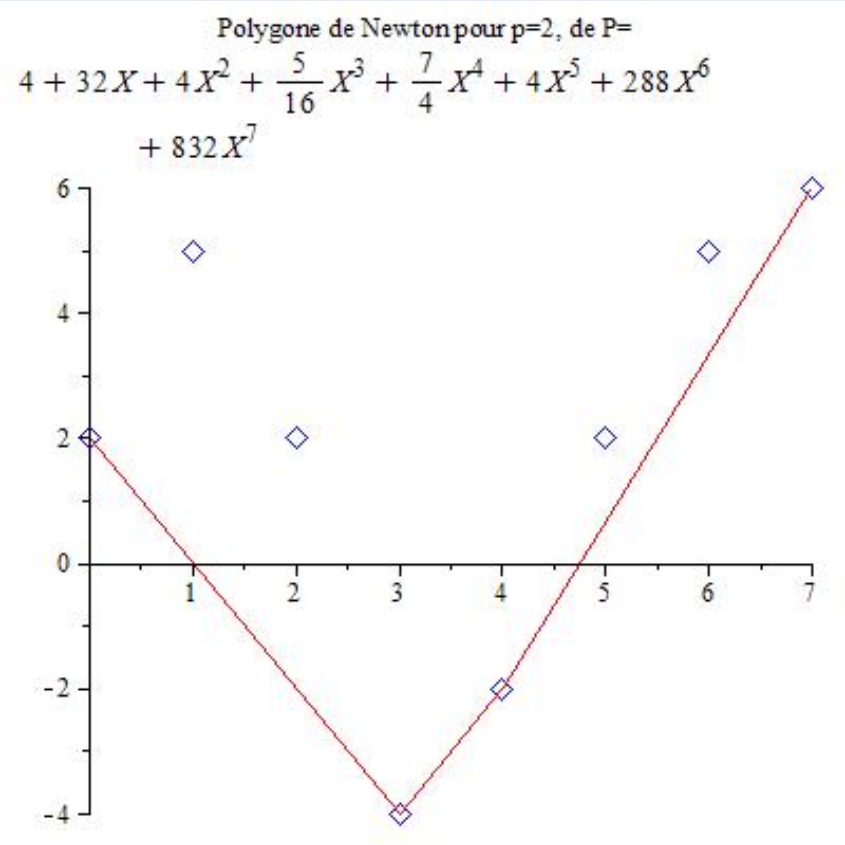


Figure: Un premier exemple

Vocabulaire et théorème fondamental des polygones de Newton

Définition

On appelle *pente du polygone de Newton* un élément de $Newt'_f([0, n])$. Si λ est une pente de $Newt_f$, on appelle *segment de pente* λ de $Newt_f$ l'ensemble $\{(x, Newt_f(x)) / Newt'_f(x) = \lambda\}$. La *longueur* de ce segment sera, par définition, la longueur de son projeté sur l'axe des abscisses.

Vocabulaire et théorème fondamental des polygones de Newton

Définition

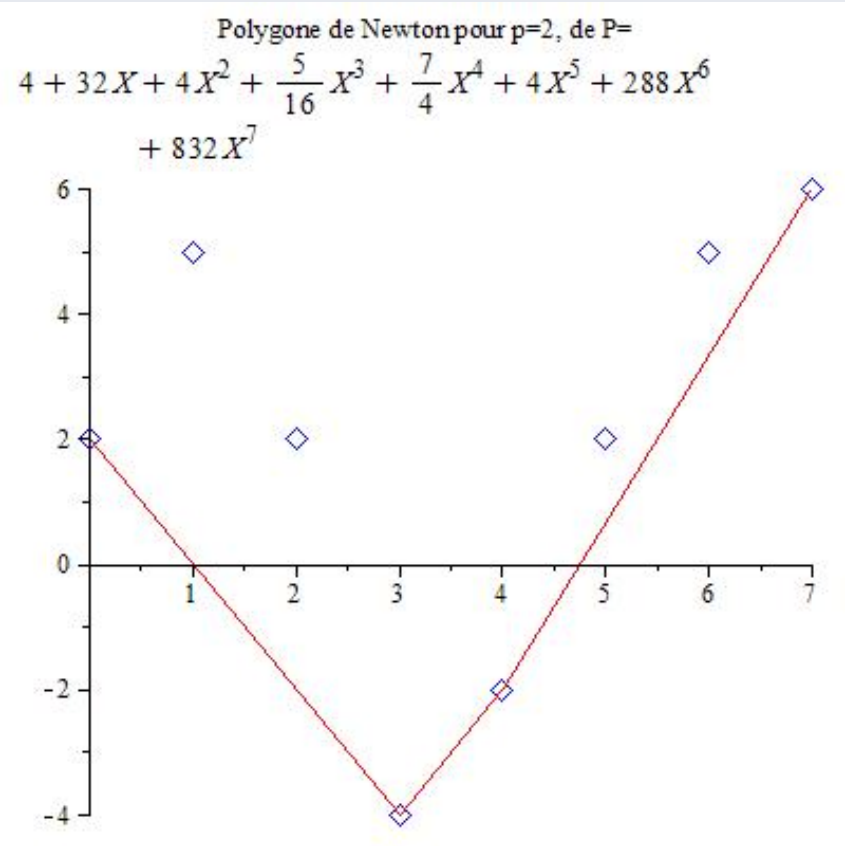
On appelle *pente du polygone de Newton* un élément de $Newt'_f([0, n])$. Si λ est une pente de $Newt_f$, on appelle *segment de pente* λ de $Newt_f$ l'ensemble $\{(x, Newt_f(x)) / Newt'_f(x) = \lambda\}$. La *longueur* de ce segment sera, par définition, la longueur de son projeté sur l'axe des abscisses.

Théorème

f a une racine de valuation λ si et seulement si $-\lambda$ est une pente de $Newt_f$. De plus, le nombre de racines de f de valuation λ, comptées avec multiplicité, est la longueur du segment de pente $-\lambda$ de $Newt_f$.

Retour à l'exemple

Valuation des racines



P a trois racines de valuation $1/2$, une de valuation -2 et trois de valuation $-8/3$.

Quelques remarques

Remarque

Du fait que si $P \in K[X]$ est irréductible, toutes ses racines ont même valuation, on obtient que si P est irréductible, alors $Newt_P$ n'a qu'une pente. La réciproque est, bien sûr, fausse (prendre par exemple $(X-1)(X-2)$ sur $\mathbb{Q}$, muni de v_5).

Quelques remarques

Remarque

Du fait que si $P \in K[X]$ est irréductible, toutes ses racines ont même valuation, on obtient que si P est irréductible, alors $Newt_P$ n'a qu'une pente. La réciproque est, bien sûr, fausse (prendre par exemple $(X-1)(X-2)$ sur $\mathbb{Q}$, muni de v_5).

Remarque

Ainsi, si le polygone de Newton de P n'est pas un segment, P n'est pas irréductible. C'est en particulier le cas dans l'exemple donné dans la Figure 1.

multiplicativité

Proposition (multiplicativité)

Si f et g sont deux polynômes, alors le polygone de Newton de fg a pour pentes celles de f et g, avec longueur la somme de celle de f et de celle de g.

Plan de l'exposé

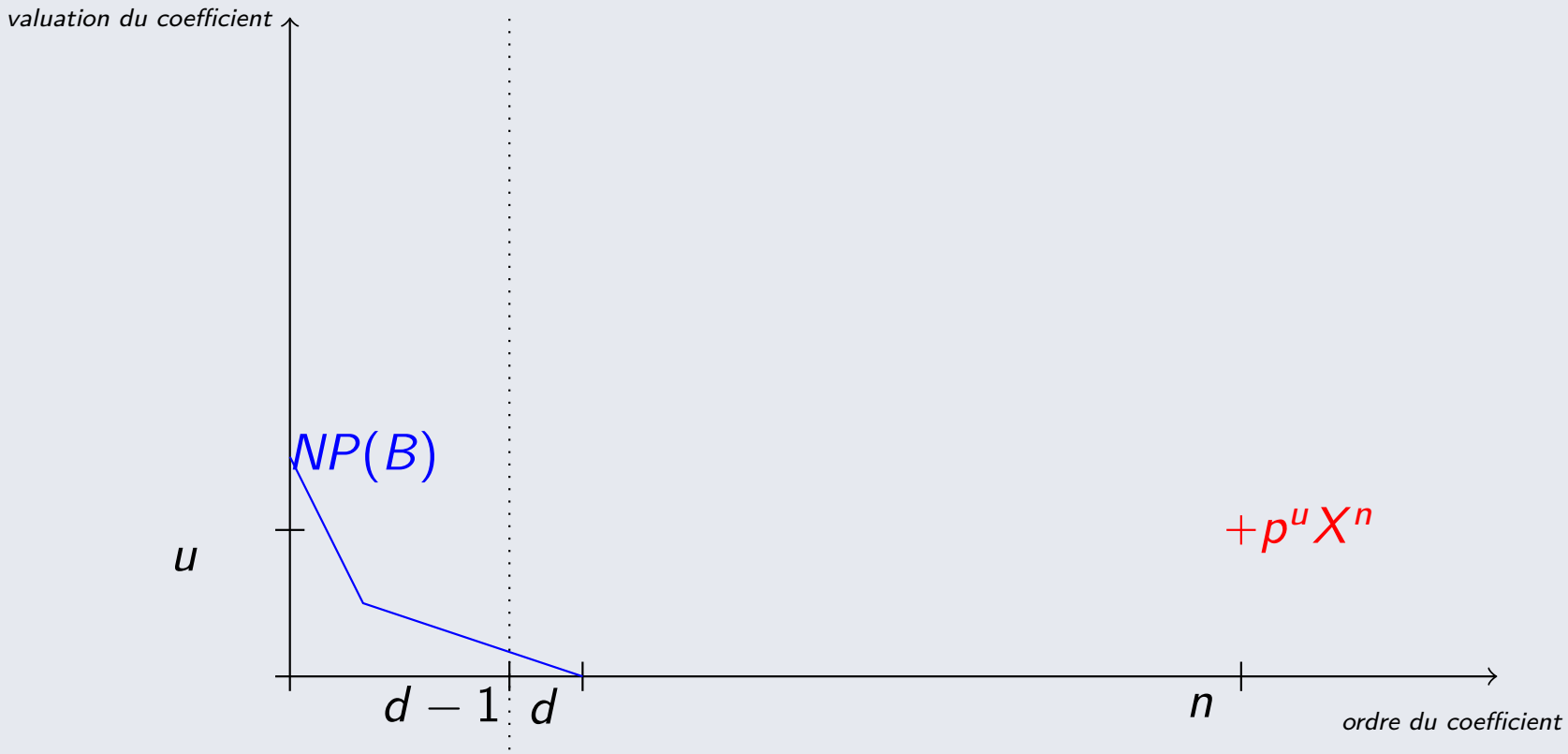
Division euclidienne et polygone de Newton

Lemme (Lemme de division)

On veut connaître les polygones de Newton du reste et du quotient de la division de A par B, deux polynômes. Considérons d'abord la division de X^n par le polynôme B, puis celle de A.

Division euclidienne et polygone de Newton

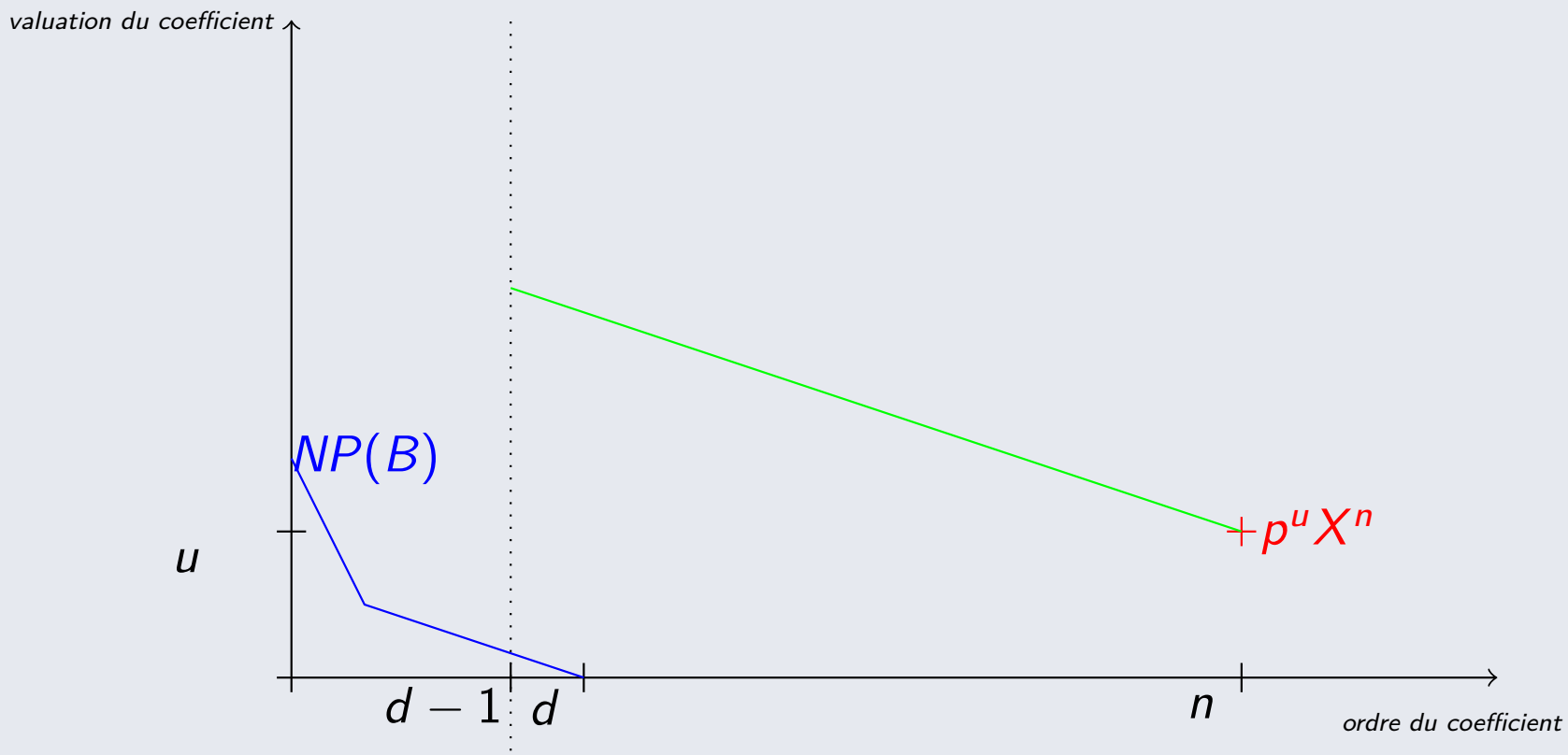
Lemme (Lemme de division)



Division de $p^u X^n$ *par* B : $p^u X^n = BQ + R$

Division euclidienne et polygone de Newton

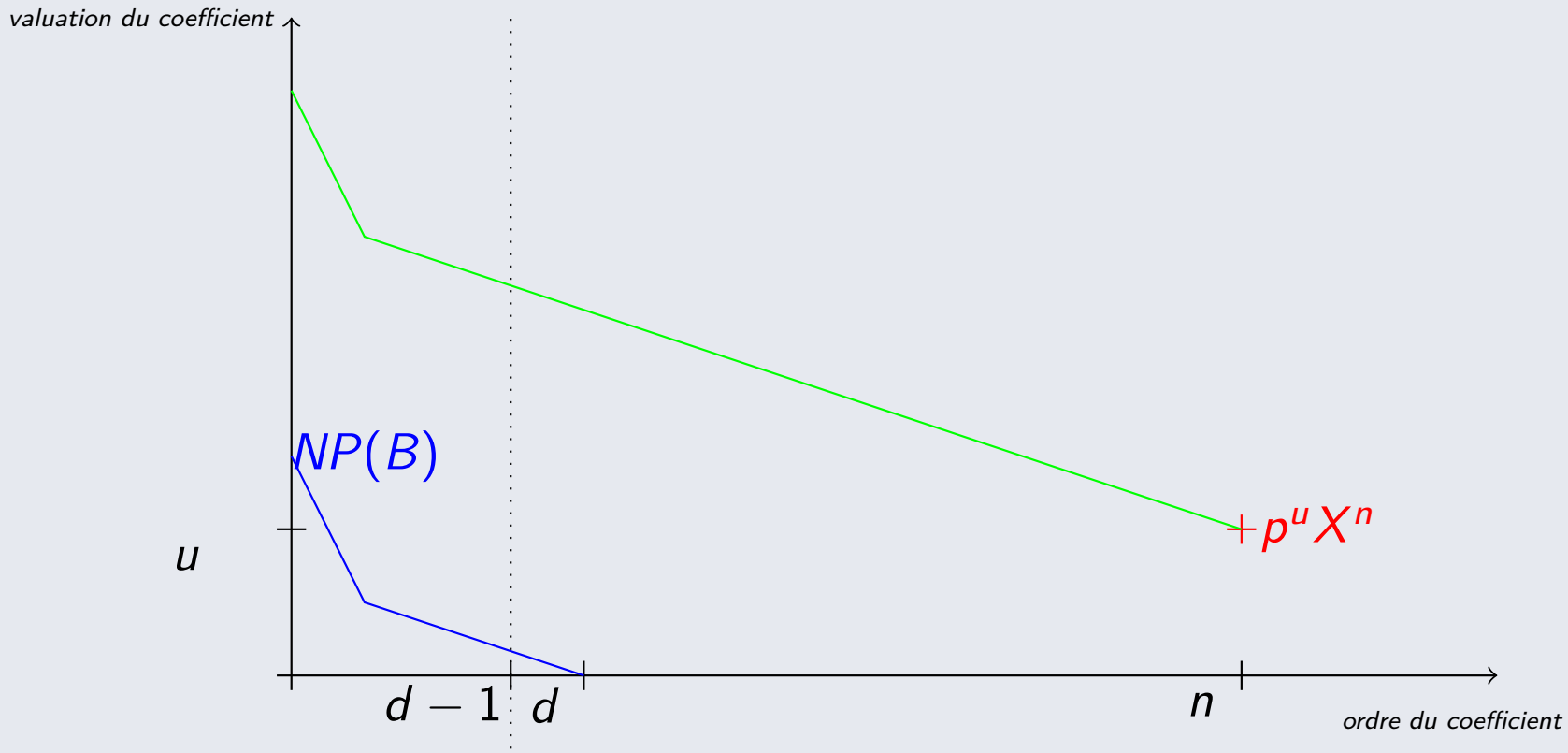
Lemme (Lemme de division)



Division de p^uX^n par B : $p^uX^n = BQ + R$

Division euclidienne et polygone de Newton

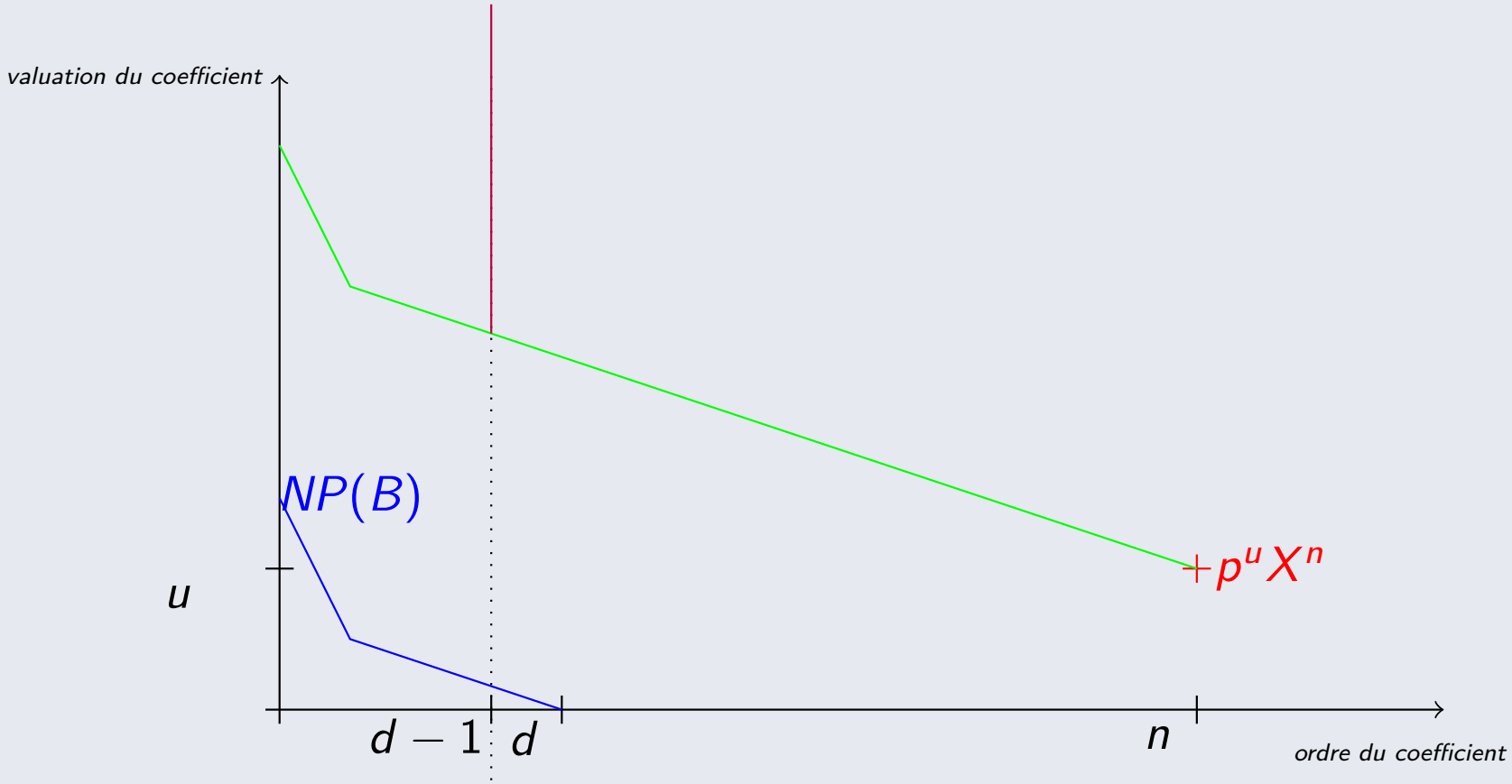
Lemme (Lemme de division)



Division de p^uX^n par B : $p^uX^n = BQ + R$

Division euclidienne et polygone de Newton

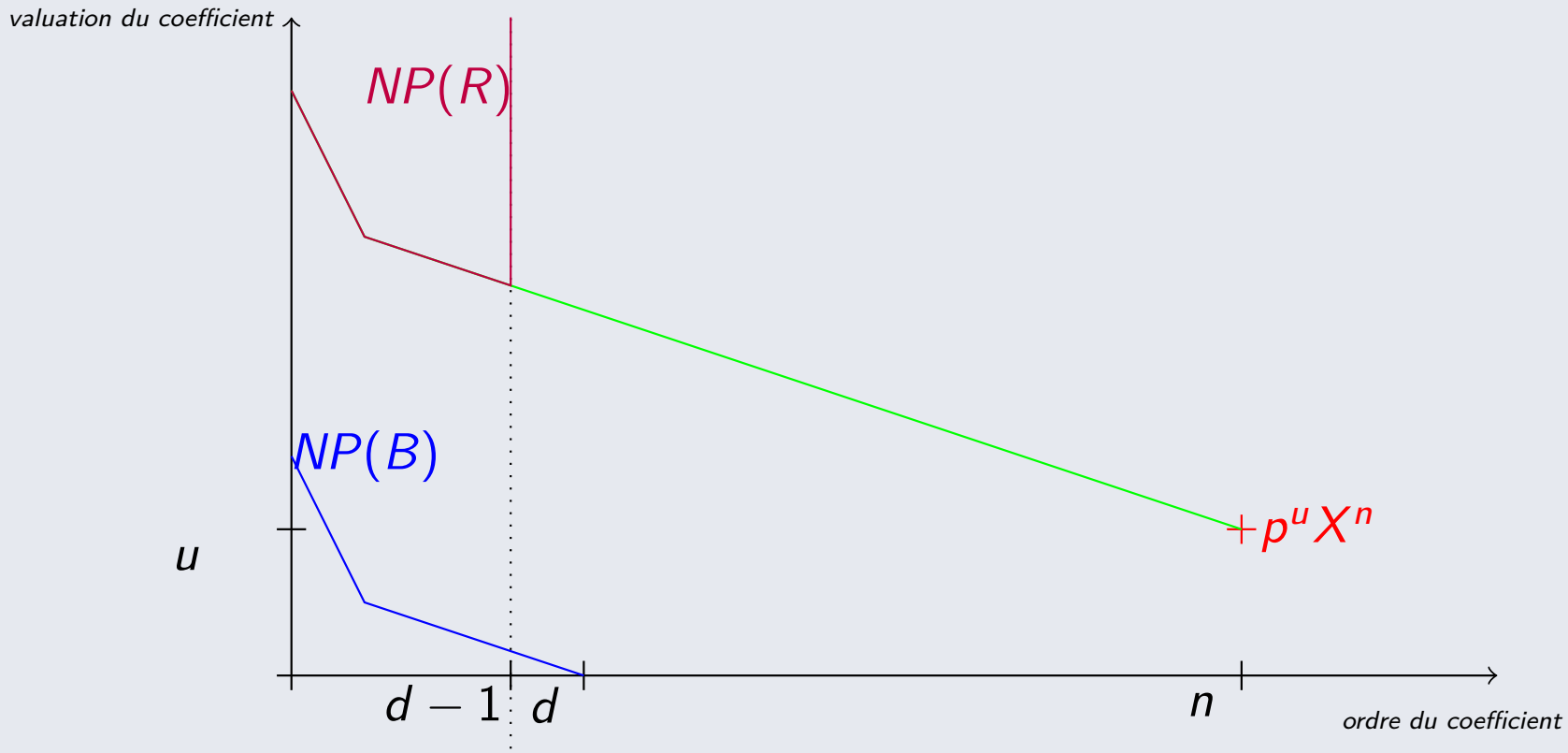
Lemme (Lemme de division)



Division de p^uX^n par B : $p^uX^n = BQ + R$

Division euclidienne et polygone de Newton

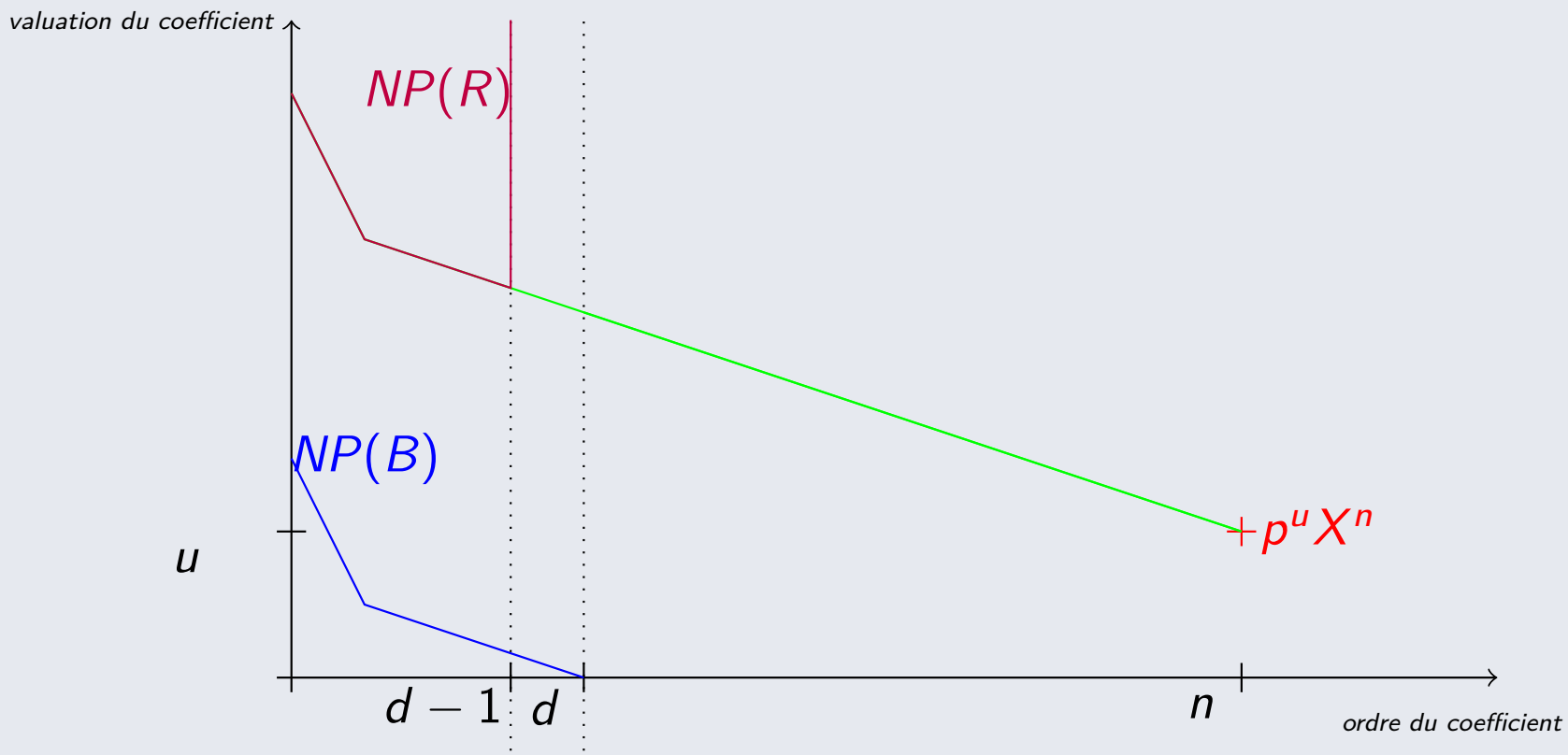
Lemme (Lemme de division)



Division de $p^u X^n$ par B : $p^u X^n = BQ + R$

Division euclidienne et polygone de Newton

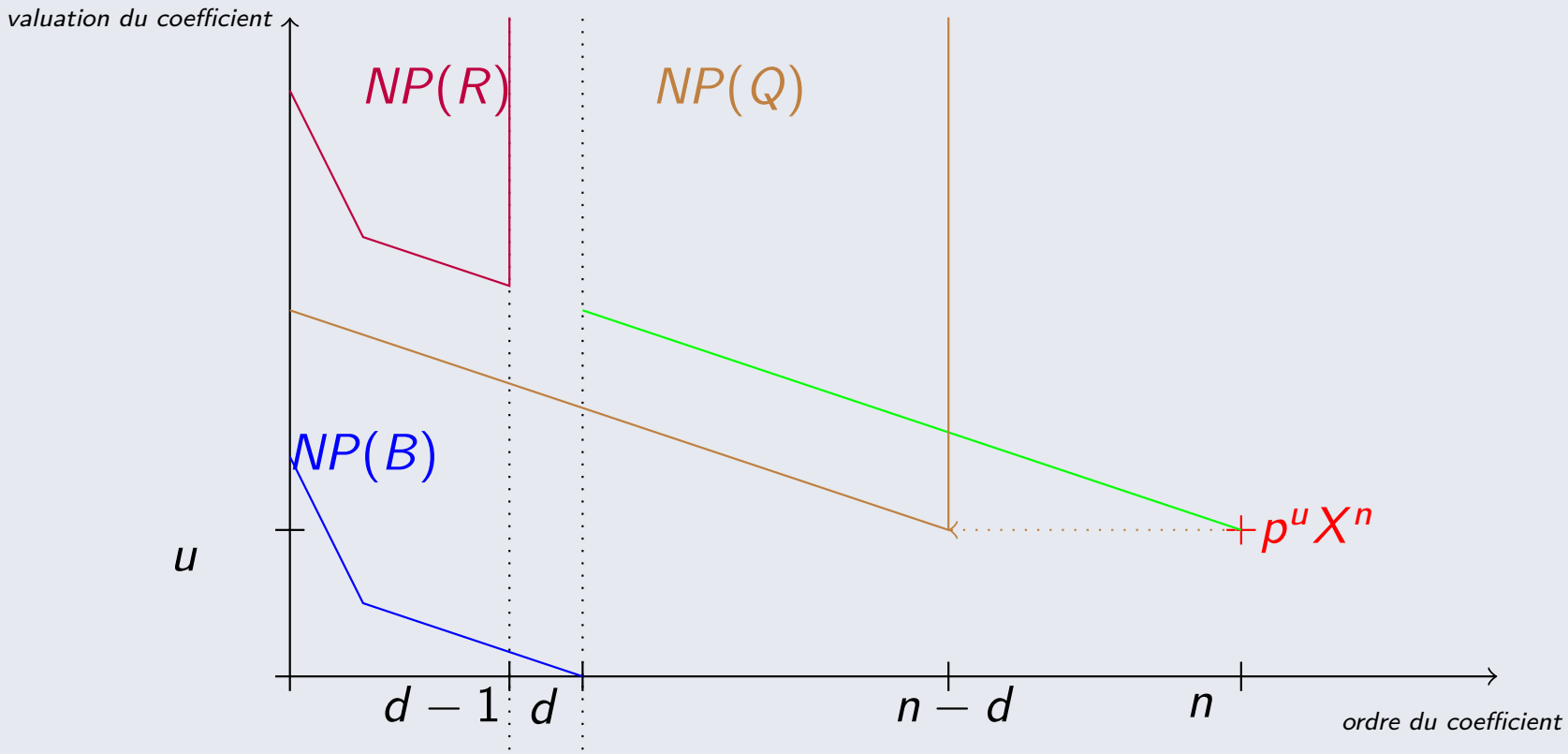
Lemme (Lemme de division)



Division de $p^u X^n$ par B : $p^u X^n = BQ + R$

Division euclidienne et polygone de Newton

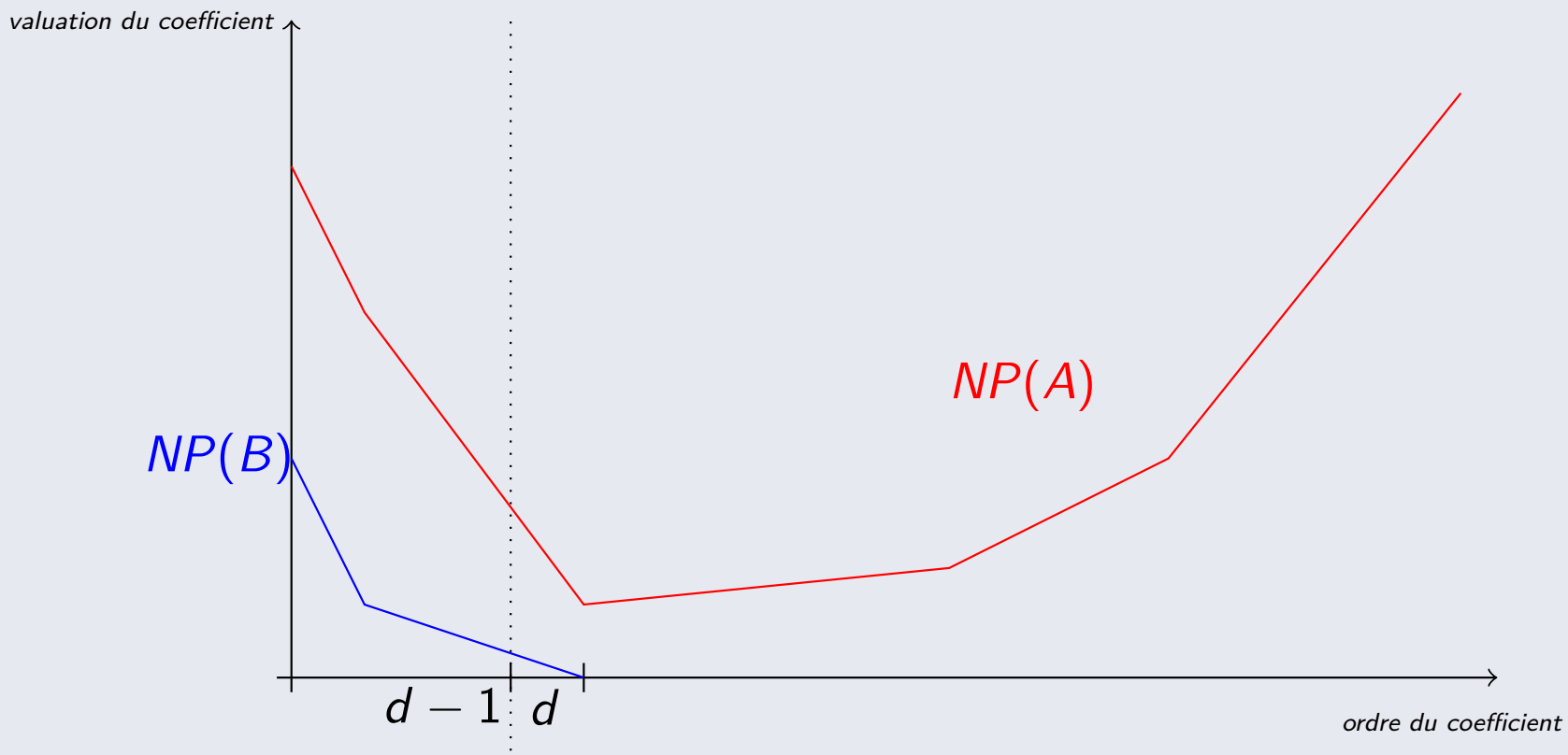
Lemme (Lemme de division)



Division de p^uX^n *par* B : $p^uX^n = BQ + R$

Division euclidienne et polygone de Newton

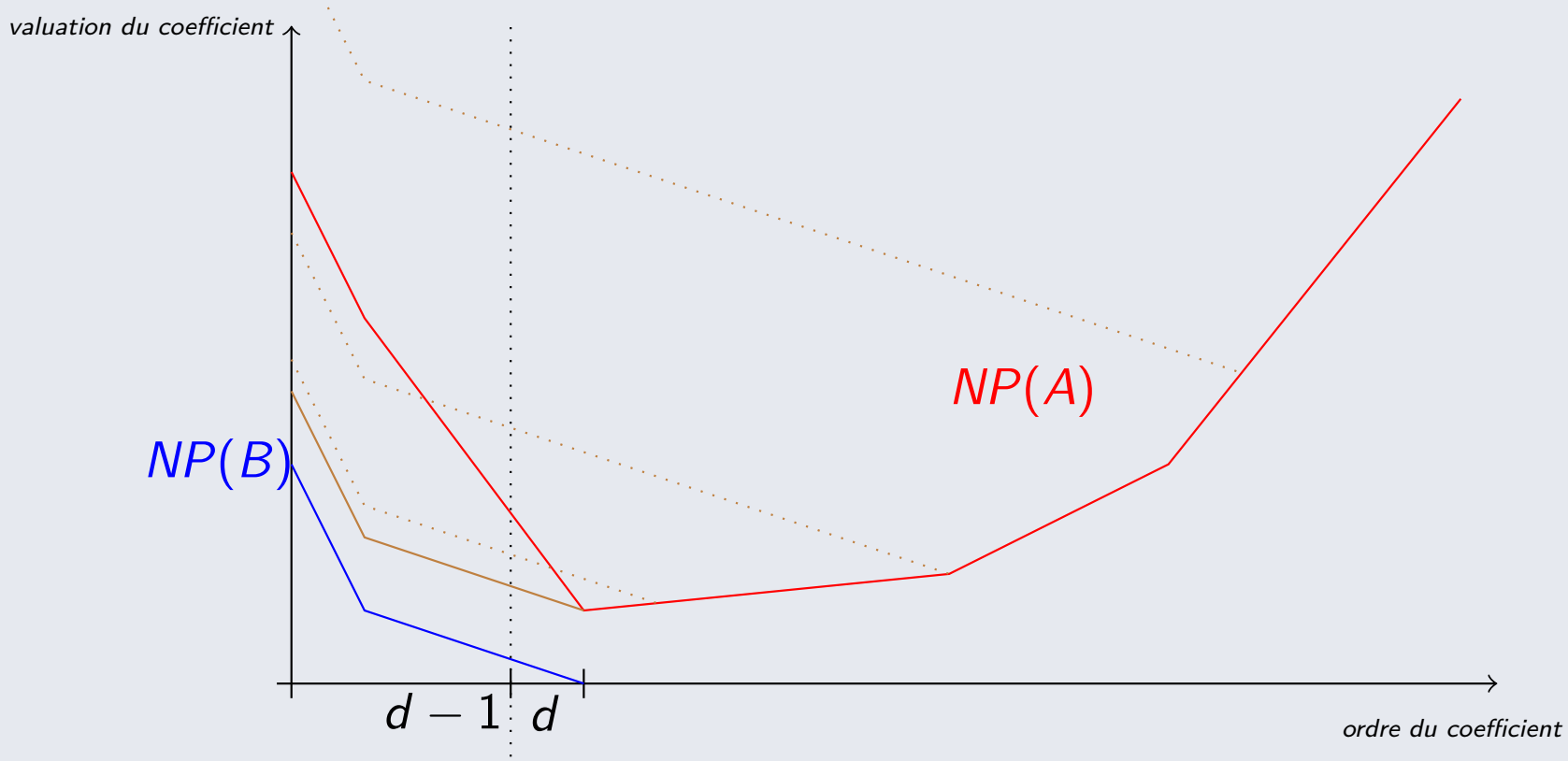
Lemme (Lemme de division)



Division de A par B : $A = BQ + R$

Division euclidienne et polygone de Newton

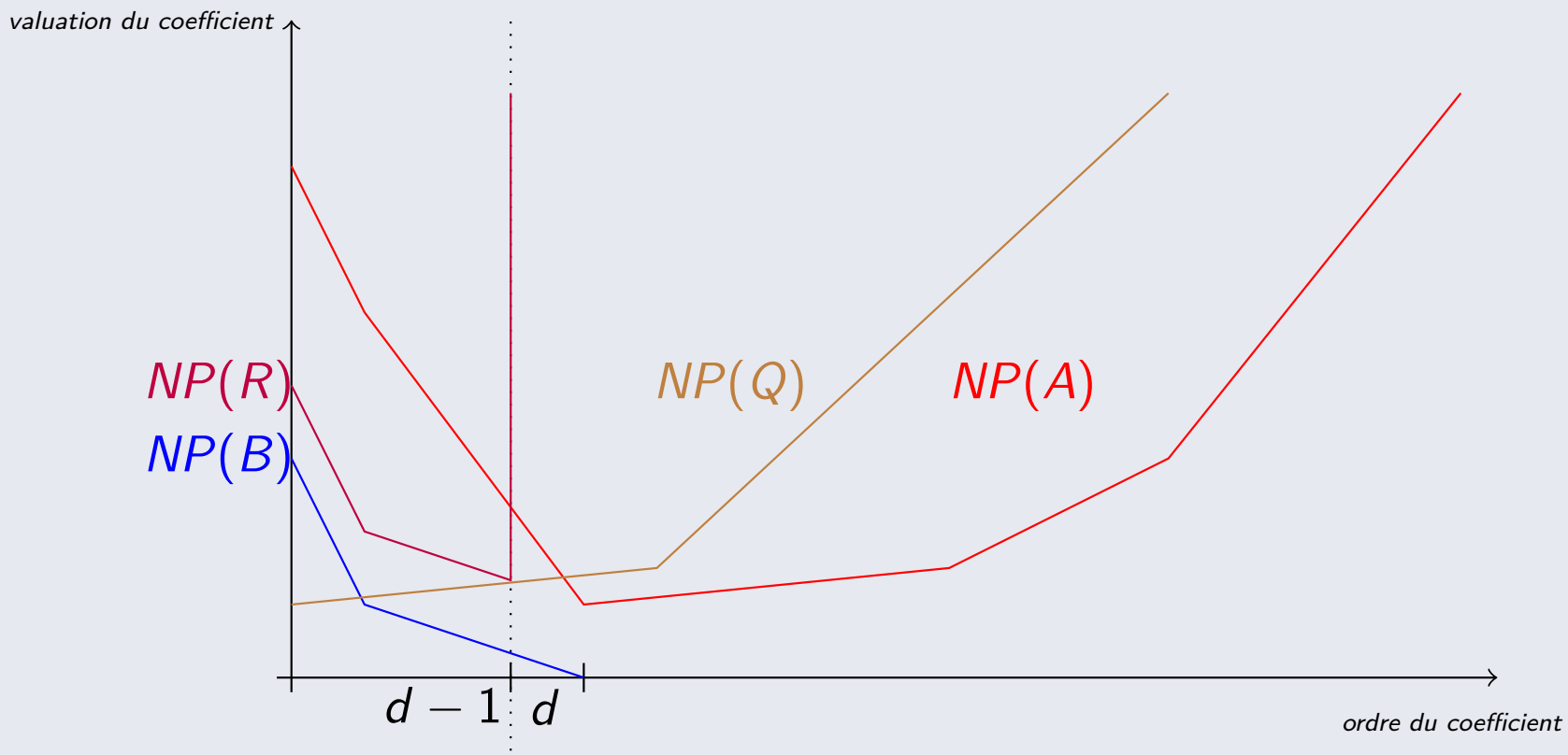
Lemme (Lemme de division)



Division de A par B : $A = BQ + R$

Division euclidienne et polygone de Newton

Lemme (Lemme de division)



Division de A par B : $A = BQ + R$

Théorème

Soit $P = \sum_{i=0}^{+\infty} a_i X^i \in \mathbb{Q}_p[X]$ *ou* $\mathbb{Q}_p[[X]]$. *Soit* d *l'abscisse d'un sommet de son polygone de Nexton. Soit* $A_0 = Lo(P, d) = \sum_{i=0}^{d} a_i X^i$,

Théorème

Soit $P = \sum_{i=0}^{+\infty} a_i X^i \in \mathbb{Q}_p[X]$ *ou* $\mathbb{Q}_p[[X]]$. *Soit* d *l'abscisse d'un sommet de son polygone de Nexton. Soit* $A_0 = Lo(P, d) = \sum_{i=0}^{d} a_i X^i$, *et définissons par récurrence*

$$A_{i+1} = A_i + (P \mod A_i).$$

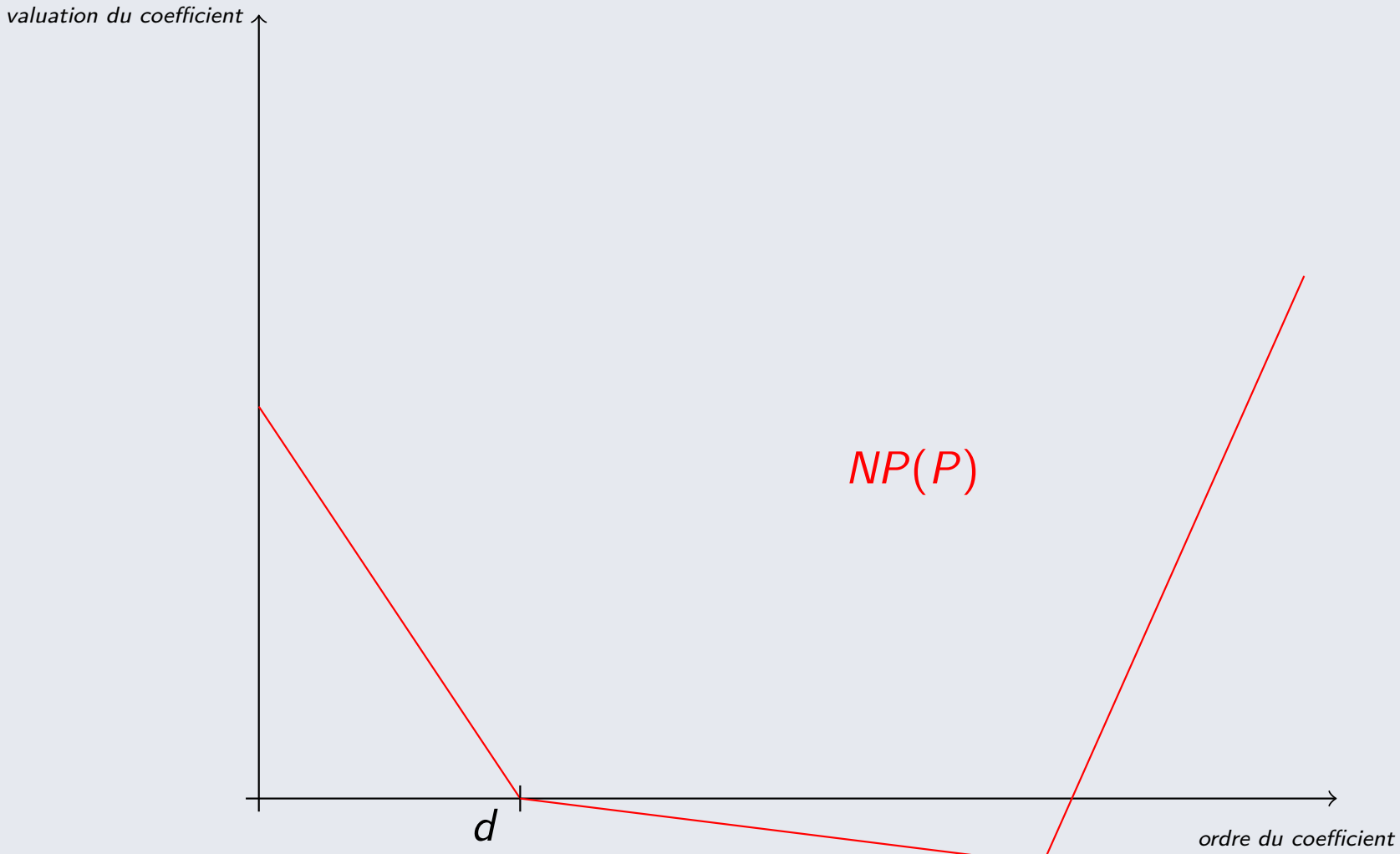
Théorème

Soit $P = \sum_{i=0}^{+\infty} a_i X^i \in \mathbb{Q}_p[X]$ *ou* $\mathbb{Q}_p[[X]]$. *Soit* d *l'abscisse d'un sommet de son polygone de Nexton. Soit* $A_0 = Lo(P, d) = \sum_{i=0}^{d} a_i X^i$, *et définissons par récurrence*

$$A_{i+1} = A_i + (P \mod A_i).$$

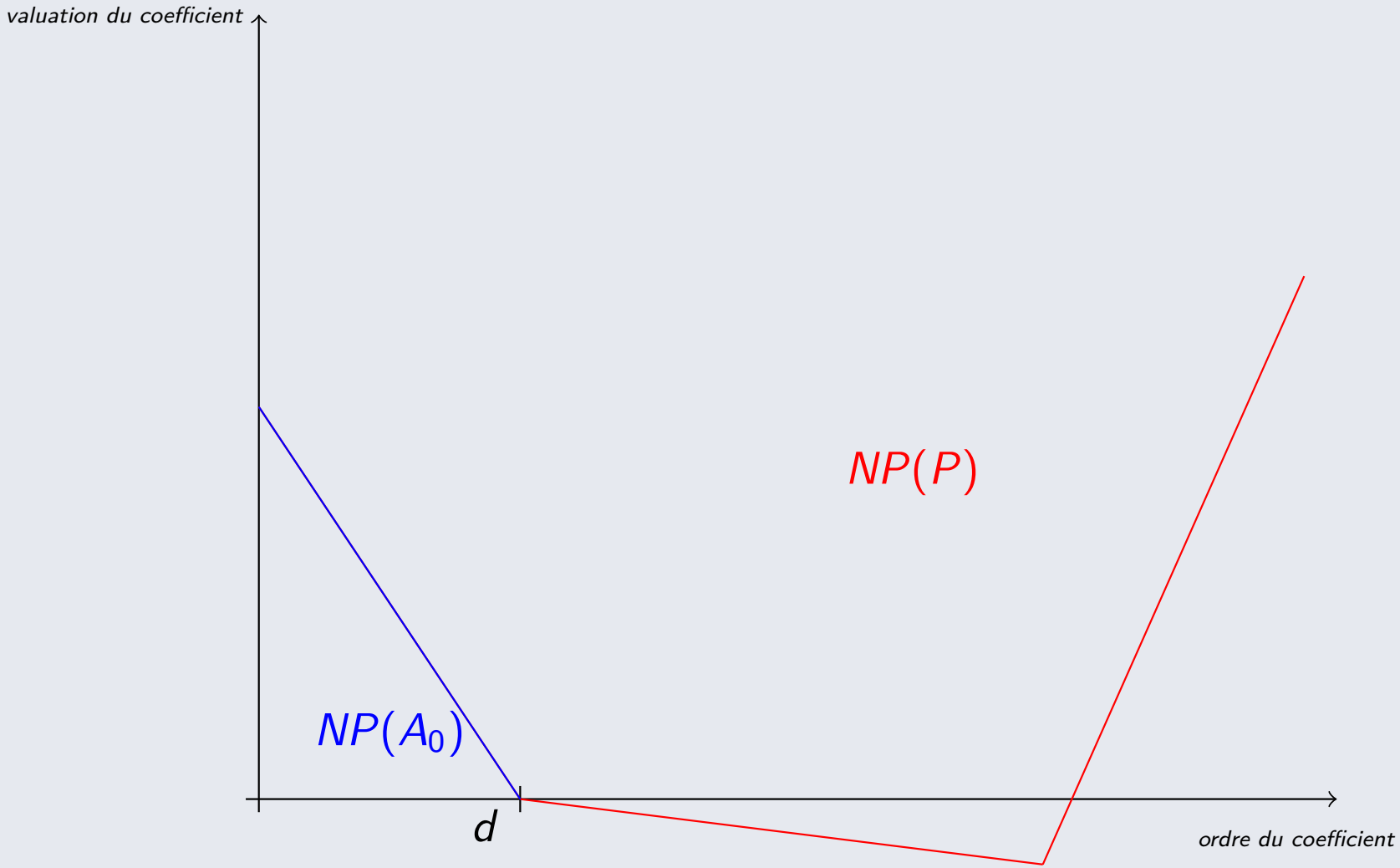
Alors la suite (A_i) *converge vers un certain polynome* A *tel que* $NP(A) = NP(Lo(P, d))$ *et* A *divise* P .

Théorème



Polygone de Newton de P

Théorème



Polygone de Newton de P

Proof.

- On peut supposer A_0 unitaire.
- On écrit la division euclidienne par A_i : $P = A_i B_{i+1} + (P \mod A_i)$.
- On pose $R_i = P - A_i B_i$.
- On alors :

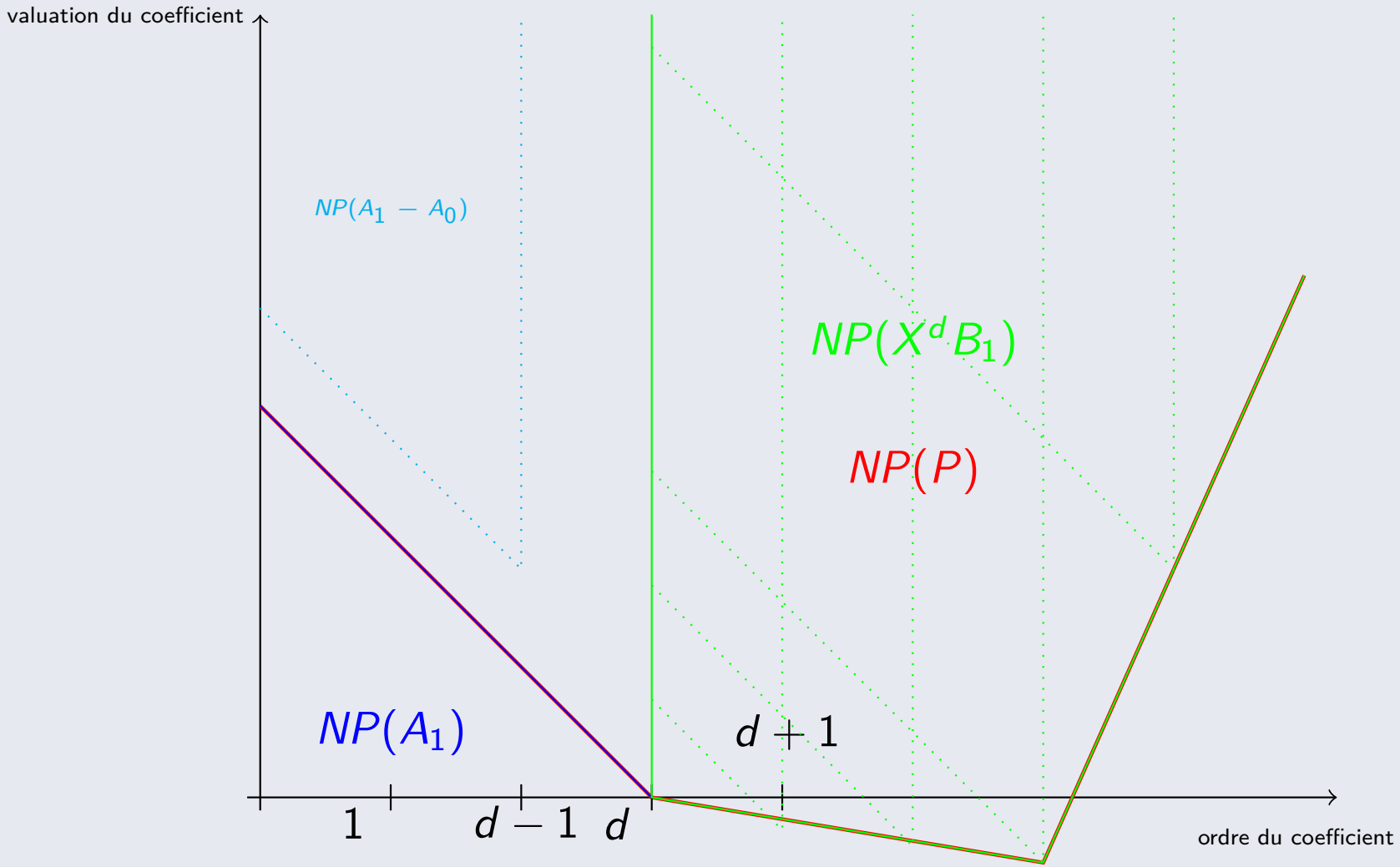
$$A_i - A_{i+1} = -(P \mod A_i) = (R_i \mod A_i).$$

- De plus :

$$R_{i+1} = (B_{i+1} - 1)(A_i - A_{i+1}).$$

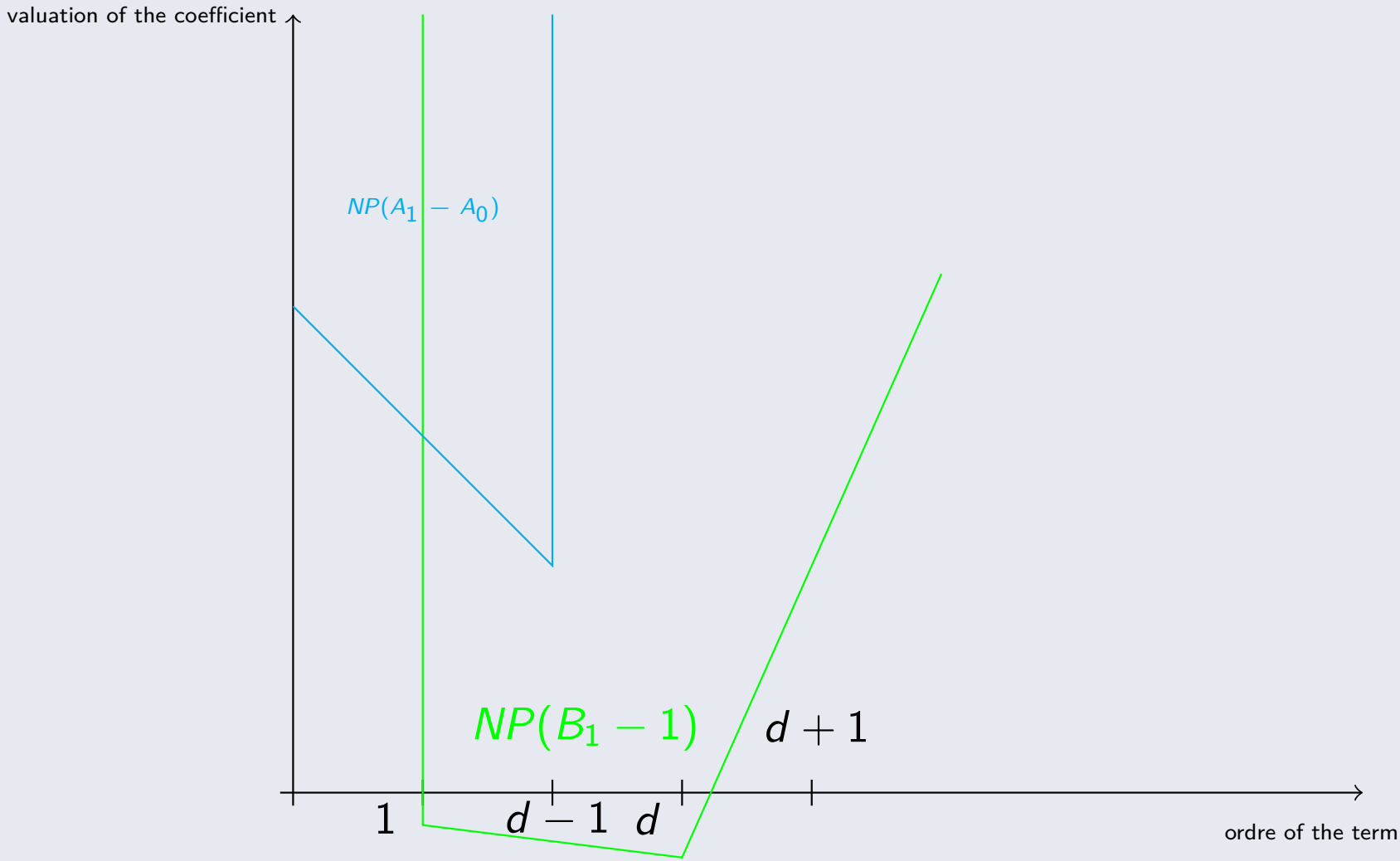
□

Proof.



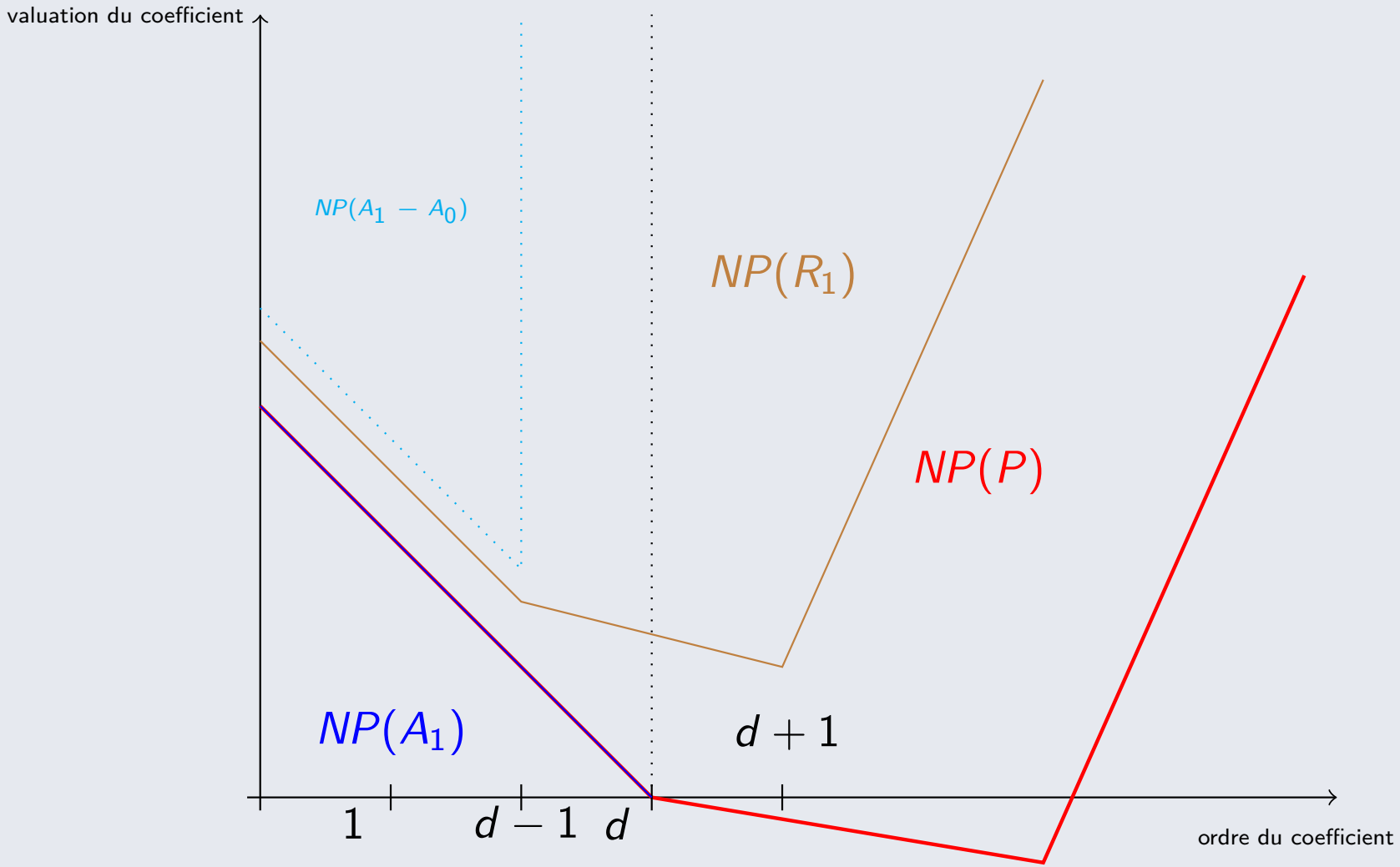
Division of P by A_0

Proof.



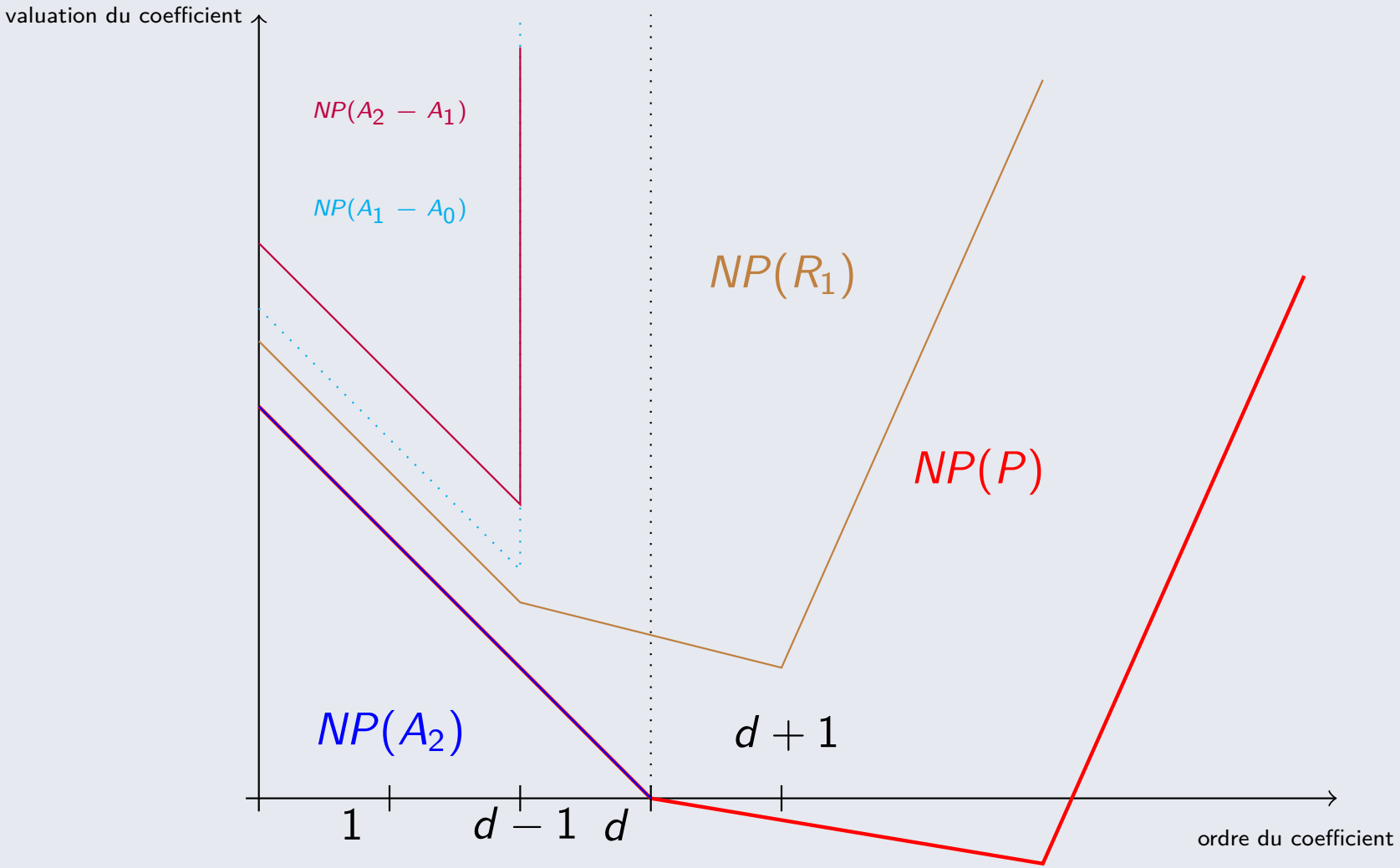
Polygones de Newton de $B_1 - 1$ et $A_1 - A_0$

Proof.



Polygone de Newton de R_1

Proof.



Polygone de Newton de $A_2 - A_1$

Proof.

Remarque

En p-adique, les notions de convergence sont plus simples :

$$A_{i+1} - A_0 = \sum_{k=1}^{i} A_{k+1} - A_k.$$

Proof.

Remarque

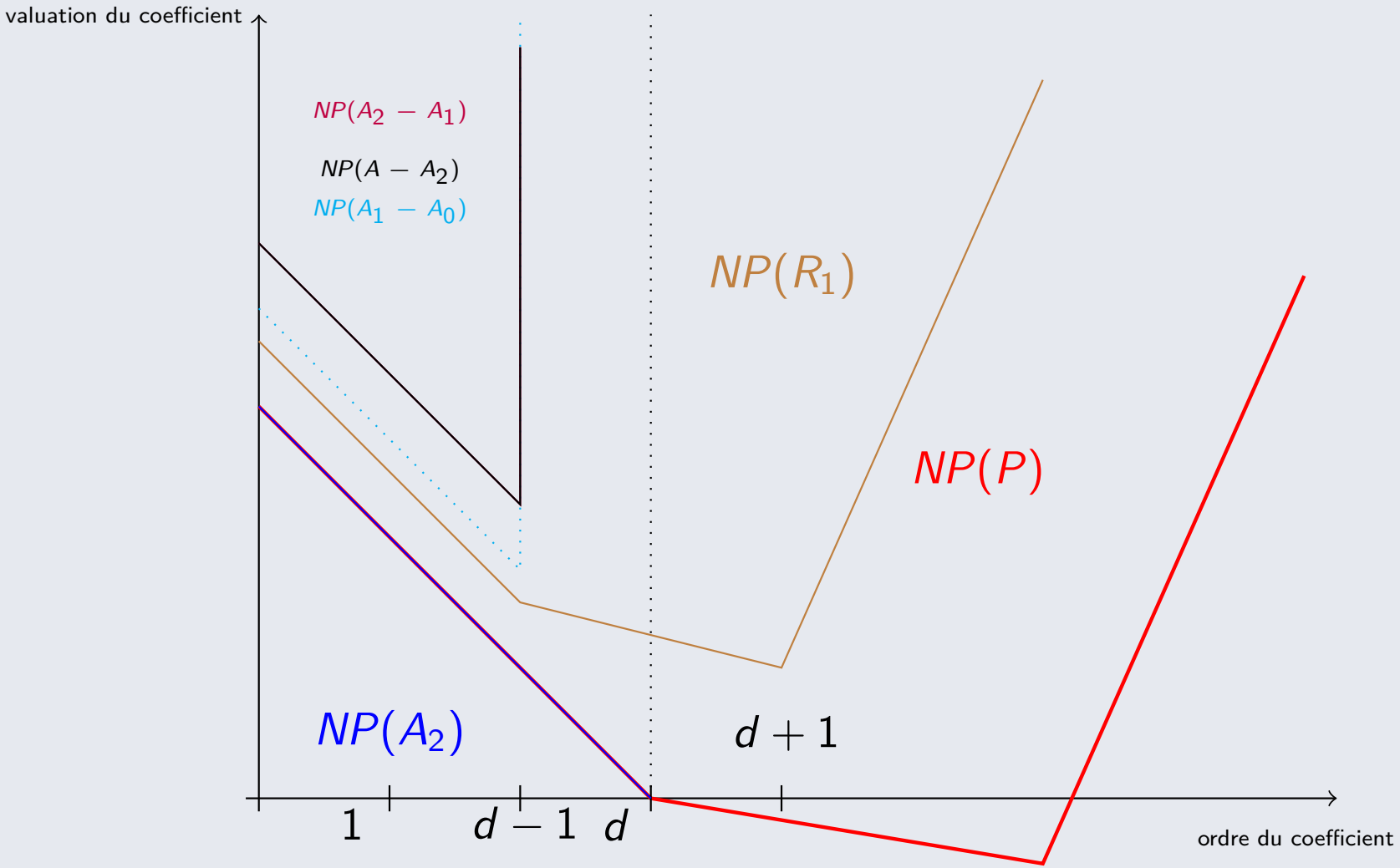
En p-adique, les notions de convergence sont plus simples :

$$A_{i+1} - A_0 = \sum_{k=1}^{i} A_{k+1} - A_k.$$

L'estimation de l'approximation aussi :

$$A - A_i = \sum_{k=i}^{+\infty} A_{k+1} - A_k.$$

Proof.



Polygone de Newton de $A - A_2$

Plan de l'exposé

Le lemme fondamental de la précision p-adique

Lemme (Caruso)

Soit f une application C^1 de $\mathbb{Q}_p^n$ dans $\mathbb{Q}_p^m$.

Le lemme fondamental de la précision p-adique

Lemme (Caruso)

Soit f une application C^1 de $\mathbb{Q}_p^n$ dans $\mathbb{Q}_p^m$.
Soit $x \in \mathbb{Q}_p^n$. On suppose que $f'(x)$ est surjective.

Le lemme fondamental de la précision p-adique

Lemme (Caruso)

Soit f une application C^1 de $\mathbb{Q}_p^n$ dans $\mathbb{Q}_p^m$.
Soit $x \in \mathbb{Q}_p^n$. On suppose que $f'(x)$ est surjective.
Alors pour toute boule assez petite $B = B(0, r)$:

$$f(x + B) = f(x) + f'(x) \cdot B.$$

Le lemme fondamental de la précision p-adique

Lemme (Caruso)

Soit f une application C^1 de $\mathbb{Q}_p^n$ dans $\mathbb{Q}_p^m$.
Soit $x \in \mathbb{Q}_p^n$. On suppose que $f'(x)$ est surjective.
Alors pour tout k assez grand :

$$f(x + p^k \mathbb{Z}_p^n) = f(x) + f'(x) \cdot p^k \mathbb{Z}_p^n.$$

Division euclidienne

Différentielle de la division euclidienne

Soit $A, B \in \mathbb{Q}_p[X]$. On veut différencier $A = BQ + R$.

Division euclidienne

Différentielle de la division euclidienne

Soit $A, B \in \mathbb{Q}_p[X]$. On veut différencier $A = BQ + R$.
On pose $A + \delta A = (B + \delta B)(Q + \delta Q) + R + \delta R$.

Division euclidienne

Différentielle de la division euclidienne

Soit $A, B \in \mathbb{Q}_p[X]$. On veut différencier $A = BQ + R$.
On pose $A + \delta A = (B + \delta B)(Q + \delta Q) + R + \delta R$.
Alors, $\delta A - Q\delta B = B\delta Q + \delta R$.

Division euclidienne

Différentielle de la division euclidienne

Soit $A, B \in \mathbb{Q}_p[X]$. On veut différencier $A = BQ + R$.
On pose $A + \delta A = (B + \delta B)(Q + \delta Q) + R + \delta R$.
Alors, $\delta A - Q\delta B = B\delta Q + \delta R$.
Ainsi, la perte de précision dans la division de $A + \delta A$ par $B + \delta B$ est de δQ sur le quotient et δR sur le reste, et ils sont donnés comme le quotient et le reste dans la division de $\delta A - Q\delta B$ par B.

Division euclidienne

Différentielle de la division euclidienne

Soit $A, B \in \mathbb{Q}_p[X]$. On veut différencier $A = BQ + R$.
On pose $A + \delta A = (B + \delta B)(Q + \delta Q) + R + \delta R$.
Alors, $\delta A - Q\delta B = B\delta Q + \delta R$.
Ainsi, la perte de précision dans la division de $A + \delta A$ par $B + \delta B$ est de δQ sur le quotient et δR sur le reste, et ils sont donnés comme le quotient et le reste dans la division de $\delta A - Q\delta B$ par B.
On peut le lire via le lemme de division concernant les polygones de Newton.

Plan de l'exposé

Présentation du problème : séries de Newton

Définition

Soit $P \in k[X]$, $P = \prod_i (X - \alpha_i)$ (sur $\overline{k}$). On définit la série de Newton de P comme

$$f_P = \sum_n \left(\sum_i \alpha_i^n \right) t^n = -\frac{(P^*)'}{P^*},$$

avec P^* le polynôme réciproque de P.

Présentation du problème : séries de Newton

Définition

Soit $P \in k[X]$, $P = \prod_i (X - \alpha_i)$ (sur $\overline{k}$). On définit la série de Newton de P comme

$$f_P = \sum_n \left(\sum_i \alpha_i^n \right) t^n = -\frac{(P^*)'}{P^*},$$

avec P^* le polynôme réciproque de P.

Remarque

Si l'on est en caractéristique nulle, connaître f_P est équivalent à connaître P. On peut même se restreindre aux premiers termes de f_P (identités de Newton).

Présentation du problème : opérations souhaitées

Proposition

On pose $P \oplus Q = \prod_{i,j}(X - \alpha_i - \beta_j)$ *et* $P \otimes Q = \prod_{i,j}(X - \alpha_i \beta_j)$.

Présentation du problème : opérations souhaitées

Proposition

On pose $P \oplus Q = \prod_{i,j}(X - \alpha_i - \beta_j)$ *et* $P \otimes Q = \prod_{i,j}(X - \alpha_i \beta_j)$.
Alors le premier correspond à la série donnée par les produits des coefficients de f_P *et* f_Q, *et le second correspond au produit de Cauchy de* f_P *et* f_Q.

Présentation du problème : opérations souhaitées

Proposition

On pose $P \oplus Q = \prod_{i,j}(X - \alpha_i - \beta_j)$ *et* $P \otimes Q = \prod_{i,j}(X - \alpha_i \beta_j)$.
Alors le premier correspond à la série donnée par les produits des coefficients de f_P *et* f_Q, *et le second correspond au produit de Cauchy de* f_P *et* f_Q.

Remarque

Comment faire sur $\mathbb{F}_p$?

Présentation du problème : opérations souhaitées

Proposition

On pose $P \oplus Q = \prod_{i,j}(X - \alpha_i - \beta_j)$ *et* $P \otimes Q = \prod_{i,j}(X - \alpha_i \beta_j)$.
Alors le premier correspond à la série donnée par les produits des coefficients de f_P *et* f_Q, *et le second correspond au produit de Cauchy de* f_P *et* f_Q.

Remarque

Comment faire sur $\mathbb{F}_p$? On relève dans $\mathbb{Z}_p$, et on réduit à la fin.

Présentation du problème : méthode de résolution

Proposition (Calcul efficace de P à partir de f_P)

On remarque que

$$(P^*)' = -f_P * P^*,$$

Présentation du problème : méthode de résolution

Proposition (Calcul efficace de P à partir de f_P)

On remarque que

$$(P^*)' = -f_P * P^*,$$

On a donc

$$y' = a(x) * y,$$

pour $a(x) = -f_P$ et $y = P^*$. Ainsi, retrouver P à partir de f_P revient à résoudre une équation différentielle (dans $\mathbb{Z}_p[[x]]$).

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Proof.

On différentie $y' = a(t)y$, en $y_0 \in \mathbb{Z}_p[[x]]$.
y_0 correspond à P^*, a à $-f_P$, tout deux dans $\mathbb{Z}_p[[x]]$

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Proof.

On différentie $y' = a(t)y$, en $y_0 \in \mathbb{Z}_p[[x]]$.

$$(y_0 + \delta y)' = (a + \delta a)(y_0 + \delta y)$$

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Proof.

On différentie $y' = a(t)y$, en $y_0 \in \mathbb{Z}_p[[x]]$.

$$(y_0 + \delta y)' = (a + \delta a)(y_0 + \delta y)$$
$$(\delta y)' = a\delta y + y_0 \delta a$$

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Proof.

On différentie $y' = a(t)y$, en $y_0 \in \mathbb{Z}_p[[x]]$.

$$(\delta y)' = a\delta y + y_0 \delta a$$

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Proof.

On différentie $y' = a(t)y$, en $y_0 \in \mathbb{Z}_p[[x]]$.

$$(\delta y)' = a\delta y + y_0 \delta a$$

Par variation de la constante,

$$\delta y = y_0 \int y_0^{-1} \times y_0 \delta a \mathrm{d}t.$$

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Proof.

Par variation de la constante,

$$\delta y = y_0 \int y_0^{-1} \times y_0 \delta a \mathrm{d}t.$$

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Proof.

Par variation de la constante,

$$\delta y = y_0 \int y_0^{-1} \times y_0 \delta a \mathrm{d}t.$$

$$\delta y = y_0 \int \delta a \mathrm{d}t.$$

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Proof.

$$\delta y = y_0 \int \delta a \mathrm{d}t.$$

Comme $\int \sum_i a_i x^i = \sum_i a_i \frac{x^{i+1}}{i+1}$, ceci donne bien une perte de précision logarithmique.

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Proof.

$$\delta y = y_0 \int \delta a \mathrm{d}t.$$

Autrement dit : si $\delta a = O(p^n)$, c'est-à-dire $\delta a = \sum_i u_i p^n t^i$, alors $\delta y = \sum_i v_i \frac{p^n}{i+1} t^i$.

Calcul de la différentielle

Théorème

La perte de précision dans le calcul du coefficient d'ordre n de P à partir de f_P est en $O(\log_p(n))$.

Proof.

$$\delta y = y_0 \int \delta a \mathrm{d}t.$$

Autrement dit : si $\delta a = O(p^n)$, c'est-à-dire $\delta a = \sum_i u_i p^n t^i$, alors $\delta y = \sum_i v_i \frac{p^n}{i+1} t^i$.

Bibliographie

Sur la précision p-adique

- Caruso, Xavier, Random matrix over a DVR and LU factorization, preprint.
- Caruso, Xavier, & Lubicz, David, Un algorithme de calcul de réseaux dans les représentations semi-stables, preprint

Sur l'algorithme de Kedlaya

- Kedlaya,Kiran, Counting points on hyperelliptic curves using Monsky-Washnitzer cohomology, Journal of the Ramanujan Mathematical Society 16 (2001),

Sur la factorisation par remontée de Hensel

- Riou, Joël, Agrégation de mathématiques, option algèbre et calcul formel : http://www.math.u-psud.fr/ riou/enseignement/2010-2011/agregation/cours.pdf