Chapter 4

Algorithme de Simon

Dans ce chapitre nous étudions un algorithme qui contient déja les ingrédients essentiels qui interviendrons dans l'algorithme de Shor. En fait ce dernier peut être vu comme l'une des généralisations possible de celui de Simon. Le problème classique de Simon est un problème avec oracle qui ne peut pas être résolu par un algorithme classique aléatoire non-exponentiel (une preuve de cette assertion existe). Par contre, comme nous le verrons, il existe un algorithme quantique (aléatoire) qui résout le problème et qui possède des complexités temporelles et spatiales polynomiales. Ainsi nous connaissons au moins un problème avec oracle pour lequel il est mathématiquement prouvé, que le calcul quantique offre une accélération exponentielle par rapport au temps de calcul classique. Cela n'implique pas que ce soit vraiment le cas aussi pour des problèmes sans oracle, bien qu'il soit assez naturel de le conjecturer (à cause notamment de l'algorithme de Shor).

4.1 Le problème de Simon

On se donne une fonction de n variables booléennes telle que

$$f : \mathbb{F}_2^n \to X, \qquad (x_1 \cdots x_n) \mapsto f(x_1 \cdots x_n) \in X$$

où X est un ensemble fini et

$$f(x_1 \cdots x_n) = f(y_1 \cdots y_n) \text{ ssi } \vec{x} = \vec{y} \text{ ou } \vec{x} - \vec{y} = \vec{a}$$

pour un vecteur $\vec{a}$ fixé. Pour nous familiariser avec une telle fonction on peut se faire l'image suivante. Il y a 2^n vecteurs d'entrée $\vec{x}$. Prenons un tel $\vec{x}$ et considérons son partenaire $\vec{x} + \vec{a}$. La fonction satisfait $f(\vec{x}) = f(\vec{x} + \vec{a})$. De plus si $\vec{x}$ et $\vec{y}$ sont différents mais ne sont pas partenaires (c'est-à-dire

$\vec{x}-\vec{y} \neq \vec{a}$) alors $f(\vec{x}) \neq f(\vec{y})$ (voir figure 4.1). Le nombre de valeurs que prend la fonction est égale au nombre de paires $(\vec{x}, \vec{x}+\vec{a})$, c'est-à-dire $\frac{2^n}{2} = 2^{n-1}$. Ainsi la cardinalité de X est nécessairement $|X| = 2^{n-1}$.

Le problème classique à résoudre est le suivant. On dispose d'un oracle :

$$\vec{x} \longrightarrow \boxed{f} \longrightarrow f(\vec{x})$$

et on doit déterminer $\vec{a}$ en posant des questions à l'oracle. Avec un algorithme classique déterministe on est assuré d'avoir la réponse en posant $2^{n-1}+1$ questions à l'oracle (au pire), et en 2 questions (au mieux). Mais de facon générique le "temps de calcul" de ce procédé est exponentiel. Considérons maintenant l'algorithme aléatoire suivant:

1. Soumettre q paires de questions $(\vec{x}_1^{(1)}, \vec{x}_2^{(1)}), \ldots, (\vec{x}_1^{(q)}, \vec{x}_2^{(q)})$ à l'oracle. On soumet les paires aléatoirement uniformément sur l'ensemble des paires (on s'assure que deux vecteurs d'une même paire sont distincts);

2. Si pour au moins une paire i la réponse est $f(\vec{x}_1^{(i)}) = f(\vec{x}_2^{(i)})$ en déduire $\vec{a} = \vec{x}_1^{(i)} - \vec{x}_2^{(i)}$ et déclarer: SUCCESS;

3. Sinon (pour toutes les paires la réponse est $f(\vec{x}_1^{(i)}) \neq f(\vec{x}_2^{(i)})$) déclarer: FAILURE;

Calculons le temps de calcul nécessaire à une probabilité de succès d'au moins $1-\epsilon$. D'apres le principe d'exclusion,

$$\Pr(SUCCESS) \leq \sum_i \Pr(f(\vec{x}_1^{(i)}) = f(\vec{x}_2^{(i)}))$$

Pour chaque terme on a $\Pr(f(\vec{x}_1^{(i)}) = f(\vec{x}_2^{(i)})) = \frac{2^{n-1}}{2^{n-1}(2^n-1)}$ puisque le nombre total de paires est $\frac{2^n(2^n-1)}{2}$ et le nombre total de paires avec égalité est 2^n. Donc

$$\Pr(SUCCESS) \leq \frac{q}{2^n-1}$$

Si on demande $\Pr(SUCCESS) \geq 1-\epsilon$ on trouve

$$q \geq (1-\epsilon)(2^n-1) \tag{4.1}$$

Tant que ϵ est fixé, il faut un nombre exponentiel de questions avec cet algorithme spécifique.

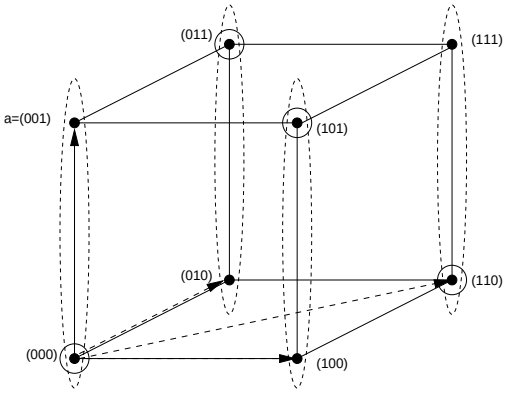

Figure 4.1: Problème de Simon dans le cas $n = 3$. On cherche le vecteur $\underline{a} = (0, 0, 1)$. Chaque ellipse représente une classe d'équivalence sur laquelle la fonction f est constante. Les points entourés sont des représentant arbitraires des classes déquivalences. Le circuit quantique retourne les vecteurs du plan perpendiculaire à $\underline{a}$

En fait il est possible de montrer que tout algorithme classique aléatoire requiert un nombre exponentiel de questions[1] et qu'essentiellement on ne pas faire mieux que ci-dessus (avec le calcul classique).

Par contre nous verrons qu'il existe un circuit quantique simple fonctionnant avec un oracle quantique, qui permet de déterminer $\underline{a}$ avec probabilité $1-\epsilon$ en un temps polynomial O ($|\ln \epsilon|$ poly(n)) (et comme d'habitude la "complexité interne de l'oracle" n'est pas prise en compte). L'oracle quantique est représenté par l'opérateur unitaire

$$U_f|\vec{x}, \vec{z}\rangle = |\vec{x}, \vec{z} + f(\vec{x})\rangle$$

Comme $f(\vec{x})$ prend 2^{n-1} valeurs il faut $n-1$ bits pour représenter ses valeurs et donc pour $\vec{z}$ aussi.

$$\begin{array}{lll} |\vec{x}\rangle \longrightarrow \bullet \longrightarrow |\vec{x}\rangle & & \text{n qubits} \\ \qquad\quad | & & \\ |\vec{z}\rangle \longrightarrow \boxed{U_f} \longrightarrow |\vec{z} + f(\vec{x})\rangle & & \text{n-1 qubits} \end{array}$$

Nous pouvons traiter en fait un problème un peu plus général par le même circuit quantique. Sa formulation est la suivante. On peut penser à

[1] Voir A. Y. Kitaev, A. H. Shen, M. N. Vyalvi "Classical and Quantum Computation" Graduate Studies in Mathematics vol 47, American Mathematical Society (2002) pp 117.

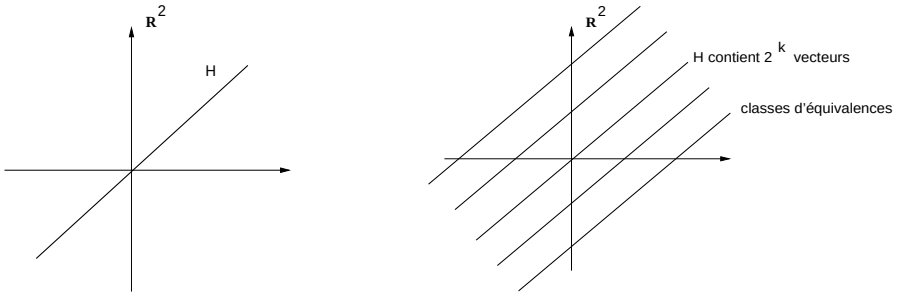

Figure 4.2: A gauche: Le sous espace vectoriel recherché H (si on remplacait $\mathbb{F}_2^n$ par $\mathbb{R}^n$). A droite: les classes d'équivalence sur lesquelles f est constante (si on remplacait $\mathbb{F}_2^n$ par $\mathbb{R}^n$).

$\mathbb{F}_2{}^n$ comme à l'espace vectoriel des vecteurs binaires à n composantes. Soit H un sous-espace vectoriel de dimension k (si on remplacait $\mathbb{F}_2^n$ par $\mathbb{R}^n$ alors H serait un hyperplan à k dimensions passant par l'origine, figure 4.2). On se donne une oracle qui sait calculer

$$f : \mathbb{F}_2^n \to X, \qquad \vec{x} \mapsto f(\vec{x})$$

tel que $f(\vec{x}) = f(\vec{y})$ ssi $\vec{x} - \vec{y} \in H$. Dans l'exemple précédent $H = \{0, \vec{a}\}$ était le sous-espace vectoriel unidimensionnel généré par $\vec{a}$. Le but est de déterminer une base de vecteurs de H en posant le moins de questions possible à l'oracle. Quelle est la cardinalité de X ? Si $\dim H = k$ alors H possède 2^k vecteurs binaires. En effet tout vecteur de H peut s'écrire $b_1\vec{h_1} + b_2\vec{h_2} + \cdots + b_k\vec{h_k}$ avec $b_i = 0, 1$. De plus $\mathbb{Z}_2^n$ n'est rien d'autre que l'ensemble de tous les "hyperplans" translatés de H. Puisque $f(\vec{x}) \neq f(\vec{y})$ si $\vec{x}$ et $\vec{y}$ n'appartiennent pas au même hyperplan, on doit avoir $|X| = \frac{2^n}{2^k} = 2^{n-k}$.

4.2 Circuit quantique pour l'algorithme de Simon

Le circuit de l'algorithme est représenté sur la figure 4.3. Sa largeur est $2n-k$ et sa profondeur est constante (égale à 3). En d'autres termes la taille du circuit est $\mathrm{O}(n)$.

Ci-dessous nous analysons en détail l'évolution de l'état quantique au cours du temps.

L'état initial à l'instant t est

$$|\Psi_{\text{in}}\rangle = \underbrace{|0\rangle \otimes \cdots \otimes |0\rangle}_{\text{n qubits}} \otimes \underbrace{|\vec{0}\rangle}_{\text{n-k qubits}}$$

À l'instant t_1 on obtient l'état

$$\begin{aligned} U(t_1,t_{\text{in}})|\Psi_{\text{in}}\rangle &= (H\otimes\cdots\otimes H)\otimes\mathbb{I}\underbrace{|0\rangle\otimes\cdots\otimes|0\rangle}_{\text{n qubits}}\otimes\underbrace{|\vec{0}\rangle}_{\text{n-k qubits}} \\ &= H|0\rangle\otimes\cdots\otimes H|0\rangle\otimes|\vec{0}\rangle \\ &= \frac{1}{2^{\frac{n}{2}}}\sum_{x_1\cdots x_n}|x_1\cdots x_n;\vec{0}\rangle \end{aligned}$$

Pour obtenir l'état à l'instant t_2 on agit avec $U(t_2,t_1)=U_f$.

$$\begin{aligned} U(t_2,t_1)U(t_1,t_{\text{in}})|\Psi_{\text{in}}\rangle &= \frac{1}{2^{\frac{n}{2}}}\sum_{x_1\cdots x_n}U_f|x_1\cdots x_n;\vec{0} \\ &= \frac{1}{2^{\frac{n}{2}}}\sum_{x_1\ldots x_n}|x_1\cdots x_n;f(x_1\cdots x_n)\rangle \end{aligned}$$

Avant de procéder plus avant, analysons la structure de cette somme. Tout vecteur $x_1\cdots x_n\in\mathbb{F}_2^n$ peut se décomposer en

$$\vec{v}+\vec{h}$$

où $\vec{v}$ caractérise l'hyperplan (il y en a 2^{n-k}) et $\vec{h}\in H$ (voir figure 4.2). On peut toujours choisir un représentant pour l'hyperplan. Dans la suite on se donne un ensemble de tels représentants $\vec{v}$, et les $\sum_{\vec{v}}$ portent sur cet ensemble de représentants. . Donc

$$\begin{aligned} U(t_2,t_1)|\Psi_{\text{in}}\rangle &= \frac{1}{2^{\frac{n}{2}}}\sum_{\vec{v}}\sum_{\vec{h}\in H}|\vec{v}+\vec{h};f(\vec{v}+\vec{h})\rangle \\ &= \frac{1}{2^{\frac{n}{2}}}\sum_{\vec{v}}\sum_{\vec{h}\in H}|\vec{v}+\vec{h};f(\vec{v})\rangle \end{aligned}$$

où on a utilisé $f(\vec{v}+\vec{h})=f(\vec{v})$ qui caractérise la fonction f. Il reste à calculer l'état à l'instant t_{fin} :

$$U(t_{\text{fin}},t_{\text{in}})|\Psi_{\text{in}}\rangle = \frac{1}{2^{\frac{n}{2}}}\sum_{\vec{v}}\sum_{\vec{h}\in H}H^{\otimes n}\otimes\mathbb{I}^{\otimes n-k}|\vec{v}+\vec{h};f(\vec{v})\rangle$$

Calculons

$$\begin{aligned} H^{\otimes n}|\vec{v}+\vec{h}\rangle &= H|v_1+h_1\rangle\otimes\cdots\otimes H|v_n+h_n\rangle \\ &= \frac{1}{2^{\frac{n}{2}}}\left(|0\rangle+(-1)^{v_1+h_1}|1\rangle\right)\otimes\cdots\otimes\left(|0\rangle+(-1)^{v_n+h_n}|1\rangle\right) \\ &= \frac{1}{2^{\frac{n}{2}}}\sum_{y_1\cdots y_n}(-1)^{\sum_{i=1}^n(v_i+h_i)y_i}|y_1\cdots y_n\rangle \end{aligned}$$

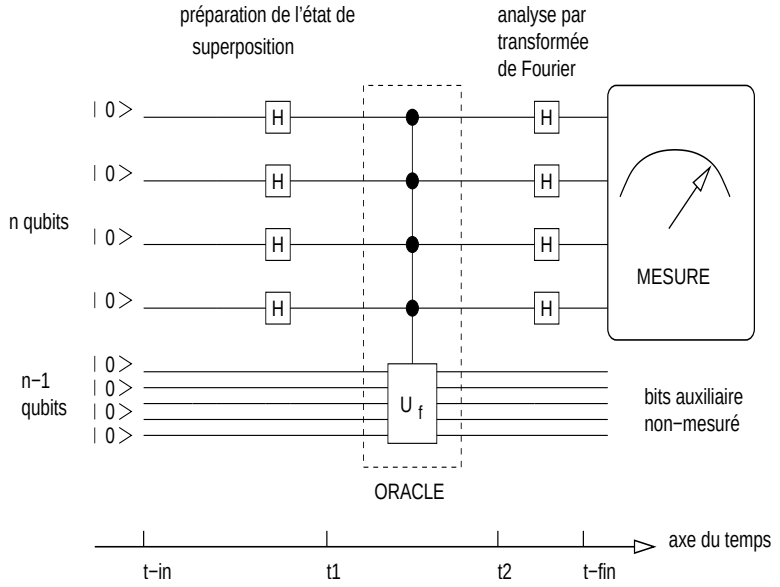

Figure 4.3: Circuit quantique pour l'algorithme de Simon

L'état à l'instant final (avant la mesure) est :

$$|\Psi_{\text{fin}}\rangle = \frac{1}{2^n}\sum_{\vec{v}}\sum_{\vec{h}\in H}\sum_{\vec{y}\in\mathbb{F}_2^n}(-1)^{(\vec{v}+\vec{h})\cdot\vec{y}}|\vec{y}\rangle\otimes|f(\vec{v})\rangle$$

$$= \frac{1}{2^n}\sum_{\vec{v}}\sum_{\vec{y}\in\mathbb{F}_2^n}(-1)^{\vec{v}\cdot\vec{y}}|\vec{y}\rangle\otimes|f(\vec{v}\rangle\left\{\sum_{\vec{h}\in H}(-1)^{\vec{h}\cdot\vec{y}}\right\}$$

Si $\vec{y}\in H^{\perp}$ (l'hyperplan perpendiculaire[2] à H) on a

$$\sum_{\vec{h}\in H}(-1)^{\vec{h}\cdot\vec{y}} = \sum_{\vec{h}\in H}1 = |H| = 2^k$$

Par contre si $\vec{y}\notin H^{\perp}$ il existe un vecteur non-nul de H, $\vec{h_0}$, tel que

$$\vec{y}\cdot\vec{h_0}\neq 0$$

Grâce au changement de variable $\vec{h} = \vec{h'} + \vec{h_0}$ on obtient :

$$\sum_{\vec{h}\in H}(-1)^{\vec{h}\cdot\vec{y}} = \sum_{\vec{h'}\in H}(-1)^{(\vec{h'}+\vec{h_0})\cdot\vec{y}}$$

$$= (-1)^{\vec{h_0}\cdot\vec{y}}\sum_{\vec{h'}\in H}(-1)^{\vec{h'}\cdot\vec{y}}$$

[2]Par définition $H^{\perp} = \{\vec{y}|\vec{y}\cdot\vec{h} = 0,\ \text{mod } 2,\ \text{tout } \vec{h}\in H\}$

Notons que $(-1)^{\vec{h_0}\cdot\vec{y}} = -1$ (car $\vec{y}\cdot\vec{h_0} \neq 0 \Rightarrow \vec{y}\cdot\vec{h_0} = 1$ dans le cas binaire). Cela implique

$$\sum_{\vec{h}\in H}(-1)^{\vec{h}\cdot\vec{y}} = 0 \quad \text{pour} \quad \vec{y} \notin H^{\perp}$$

En conclusion seuls les $\vec{y} \in H^{\perp}$ contribuent à l'état final qui peut s'écrire,

$$|\Psi_{\text{fin}}\rangle = \frac{1}{2^{n-k}} \sum_{\vec{y}\in H^{\perp}} \sum_{\vec{v}} (-1)^{\vec{v}\cdot\vec{y}} |\vec{y}\rangle \otimes |f(\vec{v})\rangle$$

À ce stade le circuit quantique produit l'état $|\Psi_{\text{fin}}\rangle$. Pour en extraire une information il faut encore faire une mesure dans la base computationnelle sur les n premiers bits. Les projecteurs associés à ce processus de mesure sont

$$P_{\vec{y}} = |\vec{y}\rangle\langle\vec{y}| \otimes \mathbb{I}^{\otimes n-k}$$

Où $\mathbb{I}$ représente le fait que les bits auxiliaires ne sont pas mesurés. L'état après la mesure est

$$\frac{P_{\vec{y}}|\Psi_{\text{fin}}\rangle}{||P_{\vec{y}}|\Psi_{\text{fin}}\rangle||}$$

avec la probabilité

$$\Pr(\vec{y}) = \langle\Psi_{\text{fin}}|P_{\vec{y}}|\Psi_{\text{fin}}\rangle$$

Si $\vec{y} \notin H^{\perp}$ on a,

$$P_y|\Psi_{\text{fin}}\rangle = 0$$

Donc

$$\Pr(\vec{y}) = 0 \quad \text{si} \quad \vec{y} \notin H^{\perp}$$

Si $\vec{y} \in H^{\perp}$ on a,

$$\begin{aligned} P_{\vec{y}}|\Psi_{\text{fin}}\rangle &= \frac{1}{2^{n-k}} \sum_{\vec{y'}\in H^{\perp}} \sum_{\vec{v}} (-1)^{\vec{v}\cdot\vec{y'}} |\vec{y}\rangle\langle\vec{y}|\vec{y'}\rangle \otimes |f(\vec{v})\rangle \\ &= |\vec{y}\rangle \otimes \frac{1}{2^{n-k}} \sum_{\vec{v}} (-1)^{\vec{v}\cdot\vec{y}} |f(\vec{v})\rangle \end{aligned}$$

Il s'ensuit que,

$$\begin{aligned} \Pr(\vec{y}) &= \frac{1}{2^{n-k}} \sum_{\vec{y}\in H^{\perp}} \sum_{\vec{v'}} (-1)^{\vec{v'}\cdot\vec{y'}} \langle\vec{y'}|\vec{y}\rangle \sum_{\vec{v}} (-1)^{\vec{v}\cdot\vec{y}} \langle f(\vec{v'})|f(\vec{v})\rangle \\ &= \begin{cases} \frac{1}{2^{n-k}} & \text{si} \quad \vec{y} \in H^{\perp} \\ 0 & \text{si} \quad \vec{y} \notin H^{\perp} \end{cases} \end{aligned}$$

Donc après une mesure on obtient un vecteur uniformément aléatoire entierement contenu dans $H^{\perp}$.

4.3 Analyse probabiliste de l'algorithme

Nous allons analyser le temps de calcul de l'algorithme suivant:

1. Initialiser le circuit avec $|0_n\rangle \otimes |0_{n-k}\rangle$ et faire une demande à l'oracle quantique.

2. Mesurer l'état final $|\Psi_{\text{fin}}\rangle$. la mesure fourni $\vec{y} \in H^{\perp}$.

3. Itérer $n-k$ fois les points 1 et 2 pour obtenir $\vec{y_1}, \ldots, \vec{y}_{n-k}$, des vecteurs de $H^{\perp}$.

4. Si $\vec{y_1}, \ldots, \vec{y_{n-k}}$ forme un ensemble de vecteurs indépendants c'est une base de $H^{\perp}$. Calculer la base duale de H, $\vec{h_1}, \ldots, \vec{h_k}$ par les méthodes usuelles d'algèbre linéaire (p.ex élimination gaussienne). Output : "Success" et donner la base de H;

5. Si $\vec{y_1}, \ldots, \vec{y_{n-k}}$ n'est pas un ensemble de vecteurs indépendants donner "Failure".

Montrons que la probabilité de succès lors d'un round $1 \to 5$ est $\geq \frac{1}{4}$. Pour cela il faut montrer que quand on tire uniformément des vecteurs aléatoires avec remise dans $H^{\perp}$ on a

$$\Pr(\vec{y_1}, \ldots, \vec{y_{n-k}} \text{ indépendants}) \geq \frac{1}{4}$$

.

Preuve : Supposons que nous avons tiré $\vec{y_1}$. Il faut que $\vec{y_1} \neq \vec{0}$ et ceci a lieu avec probabilité $= \frac{2^{n-k}-1}{2^{n-k}}$. Maintenant tirons $\vec{y_2}$. Il faut que $\vec{y_2} \neq \{\vec{0}, \vec{y_1}\}$ et ceci a lieu avec une probabilité $= \frac{2^{n-k}-2}{2^{n-k}}$. Maintenant $\vec{y_3}$. Il faut que $\vec{y_3} \notin span\{\vec{y_1}, \vec{y_2}\}$ et $span\{\vec{y_1}, \vec{y_2}\}$ contient 2^2 vecteurs. Donc $\vec{y_3} \notin span\{\vec{y_1}, \vec{y_2}\}$ avec probabilité $= \frac{2^{n-k}-2^2}{2^{n-k}}$ en itérant on trouve :

$$\begin{aligned}
&\Pr(\vec{y_1}, \ldots, \vec{y_{n-k}} \text{ indépendants}) \\
&= \frac{2^{n-k}-2^0}{2^{n-k}} \cdot \frac{2^{n-k}-2^1}{2^{n-k}} \cdot \frac{2^{n-k}-2^2}{2^{n-k}} \cdot \cdots \cdot \frac{2^{n-k}-2^{m-k-1}}{2^{n-k}} \\
&= \prod_{j=1}^{n-k} \left(1 - \frac{1}{2^j}\right) = \exp\left(\sum_{j=1}^{n-k} \ln\{1 - \frac{1}{2^j}\}\right)
\end{aligned}$$

Utilisant pour $0 \leq x \leq \frac{1}{2}$ $\ln(1-x) \geq -x(2\ln 2)$ (voir graphe 4.4) cette

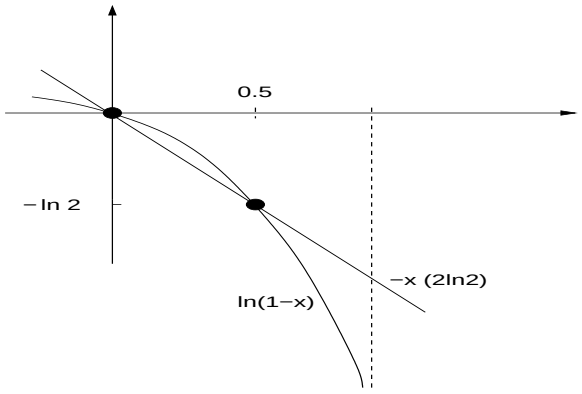

Figure 4.4: Une inegalite utile

probabilité est supérieure à :

$$\begin{aligned}
&\geq \exp\{-(2\ln 2)\sum_{j=1}^{n-k}\frac{1}{2^j}\} \\
&\geq \exp\{-(2\ln 2)\sum_{j=1}^{\infty}\frac{1}{2^j}\} \\
&= \exp(-2\ln 2) = \frac{1}{4}
\end{aligned}$$

Puisque la probabilité de succès lors d'un round est $\geq \frac{1}{4}$, la probabilité d'avoir au moins un succès lors de T rounds est

$$\begin{aligned}
\Pr(\text{au moins un succes avec } T \text{ rounds}) &= 1-(1-\Pr(\text{succes 1 round}))^T \\
&\geq 1-\left(1-\frac{1}{4}\right)^T = 1-\left(\frac{3}{4}\right)^T
\end{aligned}$$

Celle-ci est $\geq 1-\epsilon$ si et seulement si :

$$\begin{aligned}
&1-\left(\frac{3}{4}\right)^T \geq 1-\epsilon \\
&\Leftrightarrow \epsilon \leq \left(\frac{3}{4}\right)^T \\
&\Leftrightarrow T \geq \frac{|\ln\epsilon|}{|\ln\frac{3}{4}|}
\end{aligned}$$

En résumé on a une probabilité de succès $\geq 1-\epsilon$ avec $O(|\ln\epsilon|)$ rounds. Le temps de calcul de chaque round est: O (n) pour l'itération des points 1 et 2,

plus $\mathrm{O}(n^3)$ pour le calcul classique de la base duale. Donc nous avons un algorithme probabiliste polynomial avec temps de calcul $\mathrm{O}(|\ln \epsilon| n^3)$. Rappelons que la taille du circuit est $\mathrm{O}(n)$.

4.3.1 Note sur le calcul de la base duale

Soit $\vec{y_1}, \ldots, \vec{y_{n-k}}$ une base de $H^\perp$ donnée par l'algorithme. Les vecteurs de H sont orthogonaux, c'est-à-dire satisfont au système d'équations :

$$\begin{aligned} \vec{h}^T \cdot \vec{y_1} &= 0 \\ &\vdots \\ \vec{h}^T \cdot \vec{y_{n-k}} &= 0 \end{aligned}$$

Ce système possède K solutions distinctes $\vec{h_1}, \ldots, \vec{h_k}$ formant une base de H. On peut écrire le système sous forme matricielle

$$\vec{h}^T \cdot \underbrace{[\vec{y_1} \ldots \vec{y_{n-k}}]}_{\text{matrice des vecteurs colonne } n \times (n-k)} = 0$$

Donc $\vec{h}$ est dans le noyau d'une matrice de rang $n-k$. Puisque $\underbrace{\dim(\text{noyau}) + \dim(\text{image})}_{\text{rang}} = n$ on doit avoir $\dim(\text{noyau}) = k$. C'est à dire qu'il existe k solutions indépendentes $\vec{h_1} \ldots \vec{h_k}$ pour le système d'équations ci-dessus. Celles-ci peuvent être effectivement trouvées en résolvant le système par élimination gaussienne, ce qui requiert en général $\mathrm{O}(n^3)$ opérations.