HUBERT Nicolas – LAPICHE Antonin

# TP 1 : PRÉCISION DES CALCULS DANS UNE MACHINE INFORMATIQUE

## POURQUOI LE PREMIER ALGORITHME NOUS DONNE DES ERREURS IMPORTANTES ?

Le calcul de la fonction $e^x$ d'un nombre en virgule flottante se fait par une première limitation du domaine de définition de la fonction. Le résultat de calcul de cette fonction ne peut pas être supérieur au plus grand nombre représentable en double précision ni inférieur au plus petit nombre représentable en double précision de la norme IEEE 754. Les arguments qui donnent ces limites sont :

- α = log ($2^{-1023}$) = -709.0895657
- β = log ($2^{1023}$) = 709.0895657

L'exponentiel d'un nombre plus grand que β est un dépassement de capacité (OVERFLOW) alors que celui d'un nombre plus petit que α est arrondi vers zéro.

Ainsi, il faut s'intéresser au développement en série entière et à son comportement.

$$\boxed{e^x = \sum_{i=0}^{n} \frac{x^i}{i!}}$$

Si on veut connaitre $e^x$ à une erreur relative $\varepsilon$ donnée (par exemple $\varepsilon = 2^{-53}$ pour stocker le résultat dans un double) il suffira que $\frac{x^{n+1}}{(n+1)!} < \varepsilon$, on pourra donc arrêter la sommation lorsque le terme suivant sera plus petit que $\varepsilon$. Le problème est que l'on voit que plus x est grand et plus on devra itérer (plus on itère et plus on amplifie l'erreur➔retrouvé dans les résultats avec des erreurs immenses pour de grand exposants). On va donc réfléchir à éviter ce problème.

## QUELLE SOLUTION FACE À CE PROBLÈME DANS LE SECOND ALGORITHME ?

Pour calculer l'exponentielle d'un nombre x un flottant en double précision. On utilise l'identité suivante : $\boxed{e^x = 2^{\frac{x}{ln\,2}}}$

Elle provient de :

$$a^x = e^{xln(a)} \rightarrow 2^x = e^{xln\,2} \rightarrow 2^{\frac{x}{ln2}} = e^x$$

La première étape est donc de calculer l'argument réduit correspondant à la division euclidienne par $ln2$ :

$$\boxed{x = r + a \times ln2}$$

Avec r tel que :

$$\frac{-ln2}{2} < r < \frac{ln2}{2}$$

Ainsi la valeur absolue de $r$ reste bien inférieure à ln2 (définition du reste de la division euclidienne dans$\mathbb{Z}$).

Le calcul de l'exponentielle devient alors :

$$e^x = e^{r+a\times ln2} = e^r \times e^{a\times ln2} = \boxed{e^r 2^a}$$

Comme r est suffisamment petit pour ne pas contenir d'erreur, cette dernière est contenue dans le $2^a$ qui découle d'une approximation sur $ln2$ qui reste très négligeable.

Un peu + loin : Les coprocesseurs arithmétiques (FPU) possèdent des mantisses plus grandes afin de minimiser les erreurs d'arrondi de l'exponentielle et de calcul de$ln2$.

http://fr.wikipedia.org/wiki/IEEE_754

http://fr.wikipedia.org/wiki/Division_euclidienne

http://www-fourier.ujf-grenoble.fr/~parisse/mat249/mat249/node15.html

http://www.cdta.dz/sitedasm/telechargement/M1_2_Anane330.pdf