

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université de 8 Mai 1945 Guelma
Faculté des Sciences et de la Technologie
Département de génie électrotechnique et automatique



THESE

Présenté en vue de l'obtention du diplôme du **Doctorat** en Génie Electrique
Option : Réseaux électriques

**Prédiction des paramètres des décharges
couronnes dans le mélange SF₆-N₂**

Par
ASSIA GUERROUI

DIRECTEUR DE THESE :
Professeur Ahcène LEMZADMI

Université 8 Mai 1945 Guelma

DEVANT LE JURY

PRESIDENT:
Professeur Kamel BOUNAYA

Université 8 Mai 1945 Guelma

EXAMINATEURS:
Professeur Hocine LABAR
Professeur Abdelkrim MOUSSAOUI
Professeur Amar OMEIRI
Docteur El-bahi AZZAG (MC)

Université Badji Mokhtar Annaba
Université 8 Mai 1945 Guelma
Université Badji Mokhtar Annaba
Université Badji Mokhtar Annaba

ملخص

تحليل سلوك العازل الكهربائي هو من أهمية قصوى في مجال الجهد العالي. وهذا يسمح لوضع معايير الاختيار الأمثل للعزل لخفض التكاليف.

تعتبر النماذج التنبؤية لتعزيز النتائج التجريبية. حاليا نماذج التنبؤ المستخدمة تقوم أساسا على الأساليب التجريبية التي تفتقر إلى الدقة.

هدفنا هو تطوير أساليب تنبؤ أكثر ملائمة وفعالية في هذا الأطروحة نريد تقريب قانون سلوك حقيقي وتقتراح الشبكة العصبية الآلية ANN للتنبؤ بحركة الأيونات في المزيج الغازي SF₆-N₂ وذلك باستخدام البيانات التجريبية للمزيج SF₆-N₂

الشبكة المختارة هي عصيون ذو طبقة مخبأة نوع الانتشار نحو الأمام. معروف بقدرته على تقريب أي دالة خطية في البداية, نجزئ معطياتنا إلى مجموعتين, الأولى للتعليم والثانية للتحقق ثم ندرب الشبكة لكي تتعلم جيدا السلوك وأخيرا نسجل ونجرب استجابة من اجل قيم جديدة.

بالمقارنة مع البيانات التجريبية الأخطاء النسبية المتوسطة للحركة الأيونية المتوقعة هي في حدود $\pm 10\%$. وبالنظر إلى أن متوسط الأخطاء هي أقل من 10٪، لقد أثبتنا انه من المستحسن استخدام ANN المقترحة للتنبؤ بحركة الأيونات.

كلمات مفاتيح الشبكة العصبية الانتشار نحو الأمام, الضغط العالي. التنبؤ بحركة الأيونات, المزيج الغازي-SF₆

N₂

ABSTRACT

The analysis of the behavior of a dielectric insulator is of paramount importance in the field of high voltage. This allows establishing criteria for the careful choice of insulation to reduce costs.

This work is to submit dielectric insulation to extreme conditions to determine their dielectric strength and service life. Prediction models should be considered in order to consolidate the experimental results.

Currently prediction models used are mainly based on empirical methods that lack precision. Our goal is to develop more appropriate and more effective prediction methods.

This study proposes artificial neural networks (ANN) to predict the ionic mobilities in SF₆-N₂ gas mixtures. The proposed ANN consists of two input layer, four hidden layers and one output layer, which is essentially the predicted ionic mobility. In order to train the ANN, we must use experimental data for SF₆-N₂. When compared with the experimental data the average relative errors on predicted ionic mobility are found to be less than $\pm 10\%$ for training as well as for testing in all cases using the proposed ANNs. Since the average errors are less than 10%, it is recommended to use the proposed ANNs to predict the ionic mobilities.

Keywords: Neural networks modelling, feedforward, ionic mobility, High voltage, Charge carriers, corona discharge, SF₆-N₂ gas mixtures, sulphur-hexafluoride.

RESUME

L'analyse du comportement d'un isolant diélectrique est d'une importance capitale dans le domaine de la haute tension. Celle-ci permet d'établir des critères pour le choix judicieux des isolants afin de réduire les coûts.

Le présent travail consiste à soumettre des isolants diélectriques à des conditions extrêmes afin de déterminer leur tenue diélectrique et leur durée de vie.

Des modèles de prédiction sont à envisager afin de consolider les résultats expérimentaux.

Actuellement les modèles de prédiction utilisés reposent essentiellement sur des méthodes empiriques qui manquent de précision. Notre objectif est d'élaborer des méthodes de prédiction plus adaptées et plus efficaces.

Cette étude propose réseaux de neurones artificiels RNA pour prédire les mobilités ioniques dans SF₆-N₂ des mélanges de gaz. Le RNA proposé est constitué de deux couches d'entrée, quatre couches cachées et une couche de sortie, qui est essentiellement la mobilité ionique prédite. Afin de former l'RNA, en utilisant des données expérimentales pour SF₆-N₂ En comparaison avec les données expérimentales les erreurs relatives moyennes sur la mobilité ionique prédit se trouvent à moins de $\pm 10\%$ pour la formation ainsi que pour les tests dans tous les cas à l'aide des RNA proposées. Étant donné que les erreurs moyennes sont inférieures à 10%, il est recommandé d'utiliser les RNA proposées pour prédire les mobilités ioniques.

Mots-clés: Réseaux de neurones artificiels(RNA), feedforward, haute tension, Mobilité ionique, décharge couronne -Le mélange gazeux SF₆-N₂

REMERCIEMENTS

Cette thèse est réalisée au sein de Laboratoire de Génie Electrique de Guelma (LGEG).

Je tiens en premier lieu à exprimer ma gratitude très sincère à monsieur le professeur Lemzadmi Ahcène, directeur du LGEG et directeur de ma thèse pour toute sa disponibilité et pour ses conseils très utiles pour faire aboutir ce travail.

Je remercie tous les membres du Laboratoire de Génie Electrique de Guelma (LGEG) qui de près ou de loin ont contribué à l'aboutissement de ce travail.

Je voudrais aussi remercier très vivement monsieur le professeur Bounaya Kamel pour avoir bien voulu me faire l'honneur de présider le jury d'examen de ma soutenance, qu'il trouve l'expression de ma reconnaissance la plus sincère.

Je tiens à remercier particulièrement le professeur Moussaoui Abdelkrim de l'université 8 mai 1945 de Guelma. Pour sa précieuse contribution dans ce travail.

Je tiens à exprimer ma gratitude à l'égard des professeurs Omeiri Amar, Labar Hocine et Dr. Azzag El-bahi de l'université Badji Mokhtar de Annaba pour le grand plaisir qu'ils m'ont fait en acceptant de faire partie de ce jury de soutenance.

J'exprime ma très sincère reconnaissance pour tous les membres de ma famille pour les encouragements et le soutien qu'ils m'ont procurés.

ANNOTATIONS

μ	Mobilité
d	Distance inter électrode
x	Variable de distance.
q	Charge électronique (1.6×10^{-19} C).
r_p	Rayon de la pointe
r_s	Rayon du streamer.
U_c	Tension de claquage.
U_s	Tension seuil de la décharge couronne.
p	Pression exprimée en Torr
P	Pression (en bar).
N	Densité du gaz
E	Champ électrique
E_s	Champ correspondant à la charge d'espace
$(E/p)_{lim}$	Champ critique en kV/cm.bar.
u	Facteur d'utilisation du champ.
ρ	Densité du gaz.
ρ_0	Densité ionique initiale
D	Coefficient de diffusion
h	Constante de Planck

λ	Fréquence de radiation.
α	Coefficient d'ionisation
η	Coefficient d'attachement
$\bar{\alpha}$	Coefficient d'ionisation effectif
α_r	Coefficient de recombinaison électron-ion positif
N_o	Nombre d'électrons libres
γ	Coefficient secondaire d'ionisation
ϵ_0	Constante diélectrique dans le vide (8.854×10^{-12} F/m).
β	Constante qui vient d'une approximation du coefficient d'ionisation
λ	Libre parcours moyen.
\bar{v}	Vitesse thermique moyenne.
ϵ	Constante diélectrique du gaz.
α_p	Polarisabilité (en unité atomique).
ν	Nombre d'onde
V	Différence de potentiel entre deux électrodes en parallèle
μ_r	Masse réduite.
K	Critère des streamers.
e	Charge d'électron (1.6×10^{-19} C).
z	Taux du SF ₆ dans le mélange
m	Masse ionique.
M	Masse de la molécule.
D_{12}	Somme des rayons des molécules et d'ions.
a_0	Rayon de Bohr ($a=0.529177 \times 10^{-10}$ m).

V	Différence de potentiel entre deux électrodes en parallèle.
E_a	Champ harmonique.
E_c	Champ dû à la charge d'espace.
E_p	Champ électrique proche de la pointe.
I_s	Courant limité par la charge d'espace.
I	Courant moyen.
a	Distance entre l'électrode plane est la pointe $(d+r_p/2)$.
T	Temps de transit de la charge d'espace sur une ligne de champ L reliant la pointe plan
x_1 et x_2	Fractions molaires du gaz 1 et 2 dans le mélange.
N_i	Densité ionique.
w	Poids du réseau de neurones.
I_p	La jème unité cachée à l'entrée résultante
J ()	Jacobiande la performance mesurée en fonction des variables de poids et biais.
o_k	Activation de la k ^{ème} unité de sortie
RNA	Réseau de neurones
RMSE	Ecart type de l'erreur de prédiction
RP	Rétro propagation de l'erreur
s	Somme pondérée
t_k	Activation désirée de la kème unité de sortie
t	Indice d'itération
U	Entrées de système

V_{kj}	Poids d'une connexion entre la $j^{\text{ème}}$ unité cachée et la $k^{\text{ème}}$ unité de sortie
W	Vecteur poids
W_1, W_2, W_3, W_n	Coefficients synaptiques ou poids synaptiques
W_{ji}	Poids d'une connexion entre la $i^{\text{ème}}$ entrée et la $j^{\text{ème}}$ unité cachée
X	Vecteur d'entrée
X_k	l'entrée avec laquelle le poids w est associé à l'instant k .
x_1, x_2, x_3, x_n	Entrées ou excitations
i_x	La $i^{\text{ème}}$ entrée lorsque le $p^{\text{ème}}$ exemple est présenté au réseau
Y_d, y	Sortie désirée (réelle)
Δk	Différence entre la sortie attendue et la sortie effective du neurone à l'instant k
$k_j \Delta v, \Delta W_{ji}$	Termes d'adaptation des poids au cours de l'apprentissage
η	Taux d'apprentissage

TABLE DES MATIERES

ملخص.....	i
ABSTRACT.....	ii
RESUME.....	iii
REMERCIEMENTS.....	iv
ANNOTATIONS	v
TABLE DES MATIERES.....	ix
LISTE DES FIGURES	xiii
INTRODUCTION GENERALE.....	1
Chapitre 1 Généralités sur le SF ₆	4
1.1 Introduction	4
1.2 Historique sur le SF ₆	5
1.3 Propriétés du SF ₆	5
1.3.1 Propriétés physico-chimiques du SF ₆	5
1.3.2 Propriétés électriques	8
1.3.3 Le SF ₆ et l'atmosphère.....	9
1.4 Utilisation des mélanges de SF ₆	11
1.5 Propriétés de l'azote (N ₂).....	12
1.5.1 Propriétés physico-chimiques.....	12
1.6 Facteurs influençant la tenue diélectrique de l'azote.....	14
1.7 Propriétés du mélange SF ₆ -N ₂	15

1.7.1	Produits de décomposition gazeux formes dans les mélanges SF ₆ -N ₂ soumis a des décharges couronne de polarité négative	16
1.7.2	Décomposition sous l'effet de claquages de mélanges SF ₆ .N ₂	16
1.8	Conclusion	16
Bibliographie.....		18
Chapitre 2 Etude de la décharge couronne		20
2.1	Introduction	20
2.2	Processus de formation des porteurs de charge dans le SF ₆	20
2.2.1	Formation des électrons primaires.....	20
2.2.2	Ionisation, recombinaison et attachement électronique	21
2.3	Claquage diélectrique dans un gaz.....	23
2.3.1	L'avalanche électronique	23
2.3.2	Décharge de Townsend.....	24
2.3.3	Loi de Paschen.....	27
2.3.4	Déviations à la loi de Paschen	29
2.3.5	Théorie du streamer	30
2.4	Décharge couronne	33
2.4.1	Propriétés spécifiques aux décharges couronne	34
2.5	Décharge électrique dans le mélange SF ₆ -N ₂	37
2.6	La conduction électrique dans les gaz isolants.....	38
2.6.1	Théorie de la mobilité ionique de Langevin	39
2.6.2	Détermination de la mobilité par la méthode directe dite de temps de vol.....	40
2.6.3	Détermination de la mobilité à partir des courbes I-V en configuration pointe-plan	45
2.7	. Mobilité ionique dans les mélanges gazeux	50
2.8	Conclusion	51
Bibliographie.....		52
Chapitre 3 Mobilités des Porteurs de charges		54
3.1	Introduction	54

3.2	Technique expérimentale.....	54
3.2.1	Caractéristiques $I=f(U)$ et $\sqrt{I}=f(U)$ des décharges couronnes.....	55
3.2.2	Mobilité des porteurs de charges	58
3.3	Conclusion	60
	Bibliographie.....	61
	Chapitre 4 Réseaux de Neurones	62
4.1	Introduction	62
4.2	Historique	62
4.3	Les neurones formels.....	65
4.3.1	Différents types de neurones.....	66
4.4	Les réseaux de neurones formels	68
4.4.1	Les réseaux de neurones non bouclés (statiques)	68
4.4.2	Les réseaux de neurones bouclés (dynamiques)	70
4.4.3	Propriétés des Réseaux de Neurones Formels.....	71
4.5	Réseaux de neurones et régressions non linéaires.....	74
4.6	Apprentissage d'un réseau de neurones	76
4.6.1	Type d'apprentissage.....	77
4.6.3	Variantes de l'algorithme de RP.....	83
4.7	Choix des séquences d'apprentissage	87
4.7.1	Séquence des entrées de commande.....	87
4.7.2	Contraintes sur les entrées de commande :.....	87
4.7.3	Séquences d'apprentissage et estimation de la performance.....	88
4.7.4	Le surajustement	90
4.7.5	Problème des minima locaux.....	90
4.8	Etude du pouvoir de généralisation des réseaux de neurones.....	90
4.9	Validation et Tests.....	91
4.10	Conclusion.....	92
	Bibliographie.....	94

Chapitre 5 Application de réseaux de neurones pour la prédiction des mobilités dans les mélanges SF ₆ -N ₂	97
5.1 Objectif.....	97
5.2 Construction du réseau de neurones.....	97
5.2.1 Choix de l'architecture.....	99
5.2.2 Choix de la fonction d'activation.....	100
5.2.3 Algorithme d'apprentissage	100
5.3 Discussions des résultats	103
5.3.1 Apprentissage.....	103
5.4 Conclusion.....	117
Bibliographie.....	119
Conclusion générale	120
Travaux effectués.....	Erreur ! Signet non défini.

LISTE DES FIGURES

Figure 1.1: Structure moléculaire de l'hexafluorure de soufre (SF_6)	5
Figure 1.2: Tenue diélectrique du SF_6 comparé à différents matériaux en champ uniforme.....	9
Figure 1.3: Le développement de la teneur atmosphérique moyenne mondiale de SF_6	10
Figure 1.4: Evolution du champ réduit critique $(E/P)_{lim}$ (courbes construites de et tension de claquage en champ uniforme (résultats expérimentaux) de mélanges de SF_6 avec N_2 , CO_2 et air en fonction du taux de SF_6	11
Figure 1.5: Courbes de pression de vapeur pour le SF_6 , N_2 et SF_6-N_2	14
Figure 2.1: Processus d'avalanche électronique.....	24
Figure 2.2: Configuration plan-plan	25
Figure 2.3: Schéma de la relation entre deux images Tension de claquage pour différents gaz en fonction du produit p.d.....	27
Figure 2.4: Courbe de Paschen pour SF_6 en échelles logarithmiques Températures $25^\circ C$	28
Figure 2.5: Courbe de Paschen pour l'azote en échelles logarithmiques Températures $25^\circ C$	29
Figure 2.6 : Courbe de Paschen pour l'azote en échelles logarithmiques Températures $25^\circ C$	30
Figure 2.7: Mécanisme de développement du streamer proposé par Townsend.....	31
Figure 2.8: Mécanisme de propagation du streamer par photoionisation.....	32
Figure 2.9: Caractéristique courant-tension d'une décharge couronne dans l'air à pression atmosphérique (configuration pointe-plan, distance interélectrode= 13 mm)	34
Figure 2.10: (a) Lignes isopotential avant le passage du streamer	36
Figure 2.11: Topographie de la décharge couronne en configuration pointe plan	37
Figure 2.12: Comparaison entre les résultats expérimentaux de ce travail, les résultats trouvés par Malik et Quersh et ceux déterminés en utilisant Kline et al modèle pour les pur SF_6	38
Figure 2.13: Géométrie d'une configuration hyperboloïde-plan	45
Figure 2.14: 1-Distribution du champ électrique pour une conduction sans charge d'espace	47
Figure 2.15: Distribution de la densité de courant J proposé par Warburg.....	48

Figure 3.1:Dispositif de mesure des courant-tension ($I=f(U)$) pour la détermination des mobilités et des tensions seuils.....	55
Figure 3.2:Caractéristiques $I(V)$ pour le SF_6 à 100 % à différentes pressions, en polarité négative	56
Figure 3.3:Caractéristiques $\sqrt{I}=f(U)$ pour le SF_6 à 100 % pour différentes pressions, en polarité négative.....	56
Figure 3.4: Caractéristiques $I(V)$ pour le 10% SF_6 / 90% N_2 à différentes pressions, en polarité négative.....	57
Figure 3.5:Caractéristiques $\sqrt{I}=f(U)$ pour le 10% SF_6 / 90% N_2 à différentes pressions, en polarité négative.	57
Figure 3.6:Variation de la mobilité en fonction de la densité pour différente pressions.	59
Figure 3.7: Mobilités en fonction du pourcentage de l'azote N_2 dans le mélange SF_6-N_2	59
Figure 3.8:Influence du rayon de courbure de la pointe sur les mobilités des porteurs de charges, pour le mélange SF_6-N_2 à 10% de SF_6 en polarité négative	60
Figure 4.1: Neurone formel.....	65
Figure 4.2: Différentes fonctions d'activation.....	67
Figure 4.3: Perceptron multicouche à une couche cachée	69
Figure 4.4: Forme canonique d'un réseau de neurones bouclé.....	71
Figure 4.5: Définition des notations pour un PMC	79
Figure 4.6: Organigramme de conception d'un réseau de neurones.....	92
Figure 5.1:Schéma de la structure neuronale utilisée pour représenter la Mobilité des porteurs de charges	98
Figure 5.2:Type de fonctions d'actiation utilisées.....	100
Figure 5.3:Schéma bloc de l'algorithme de prédiction	101
Figure 5.4:Evolution de l'erreur.....	103
Figure 5.5:Mobilités ioniques mesurées normalisées comparés aux résultats obtenus par la prédiction des réseaux neurones à 4 bars en polarité négative	105
Figure 5.6:Evolution du taux de l'erreur de prédiction [%]	105
Figure 5.7:Mobilité ionique mesurée comparé au résultat obtenu par la prédiction des réseaux neurones à 4 bars en polarité négative.....	106
Figure 5.8:Mobilités prédites en fonction du pourcentage de l'azote N_2 dans le mélange SF_6-N_2	107
Figure 5.9:présentent la courbe normalisée des mobilités ioniques prédites par les réseaux neurones pour le mélange SF_6-N_2 à 5% de SF_6	108
Figure 5.10:Evolution du taux de l'erreur de prédiction à 5% SF_6	109

Figure 5.11: Mobilités ioniques mesurées et prédites pour un mélange à 5% de SF_6 en polarité négative.....	109
Figure 5.12: Mobilités ioniques mesurées et prédites (normalisées) pour un mélange à 10% de SF_6 en polarité négative.	110
Figure 5.13: Evolution du taux de l'erreur de prédiction à 10 % SF_6	110
Figure 5.13: Mobilités ioniques mesurées et prédites pour un mélange à 10% de SF_6 en polarité négative.....	111
Figure 5.15: Courbes normalisées des mobilités ioniques prédites par les réseaux neurones pour le mélange SF_6-N_2 à 25% de SF_6 en polarité négative.....	112
Figure 5.16: Evolution du taux de l'erreur de prédiction à 25% SF_6	112
Figure 5.17: Mobilités ioniques mesurées et prédites pour un mélange à 25 % de SF_6 en polarité négative.....	113
Figure 5.18: Mobilités ioniques normalisées prédites par les réseaux neurones pour le mélange SF_6-N_2 à 100% de SF_6	114
Figure 5.19: Evolution du taux de l'erreur de prédiction à 100% SF_6	114
Figure 5.20: Mobilités ioniques mesurées et prédites pour un mélange à 100% de SF_6 en polarité négative....	115
Figure 5.21: Courbes des mobilités des porteurs de charges déterminées par les différents modèles et celles déterminées Langevin comparées et celles mesurées par W. F. Schmidt aux résultats obtenus par la prédiction des réseaux neurones pour le SF_6 en polarité négative	116
Figure 5.21: Variation de la mobilité en fonction de la densité pour différentes pressions.....	117

INTRODUCTION GENERALE

L'hexafluorure de soufre est un gaz inerte, non agressif, incolore, inodore et pratiquement dénué de toxicité. Cependant, à concentration élevée dans l'air il agit comme asphyxiant par déplacement de l'oxygène.

Les utilisations les plus importantes de l'hexafluorure de soufre ont lieu dans le domaine de l'électricité étant donné sa résistance électrique considérable aux tensions élevées :

- transformateurs,
- condensateurs,
- appareils à haute fréquence,
- accélérateurs de particules
- relais,
- disjoncteurs.

La majorité des chercheurs dans le domaine du SF₆ le considèrent comme le gaz le plus adapté du point de vue rigidité diélectrique dans les systèmes à haute tension. Devant les problèmes issus de la décomposition et son effet négatif sur l'atmosphère les recherches se sont orientées vers des gaz de substitution.

Les essais sur les mélanges perfluocarbons s'avèrent non concluant du point de vue température et pollution. Les mélanges du SF₆ avec l'Argon ont été moins étudiés et les résultats obtenus sont insuffisants pour avoir une appréciation fiable.

Le mélange SF₆- N₂ a été le sujet le plus traité théoriquement et expérimentalement.

Néanmoins des études supplémentaires sont nécessaires pour expliquer le comportement de ce mélange dans les pressions très hautes ainsi qu'avec des températures plus au moins élevées.

D'après les travaux précédents sur le mélange SF₆-N₂ la tension de rupture augmente rapidement avec des faibles concentrations de SF₆. L'idée est d'arriver à avoir le même comportement diélectrique du SF₆ à faibles concentrations et à des pressions optimales.

L'étude de la conduction électrique de l'hexafluorure de soufre SF₆ et le mélange SF₆-N₂ en géométrie pointe-plan a été effectuée en fonction de la densité et de la polarité de la pointe. Le principal intérêt de cette configuration est de générer des champs électriques très importants (plusieurs MV/cm) pour des tensions modérées. Par ailleurs, la zone de création des porteurs de charge étant localisée au voisinage immédiat de la pointe, elle se trouve séparée de la zone de transport. Enfin l'étude de la conduction en champ intense peut être conduite sans craindre d'aboutir au claquage comme c'est le cas en géométrie plan-plan (E étant uniforme entre les deux électrodes).

Le présent travail consiste essentiellement à la détermination des mobilités des porteurs de charges du SF₆ et du mélange SF₆-N₂ en utilisant les techniques des réseaux neurones. Les données expérimentales utilisées ont été obtenues précédemment au laboratoire d'électrostatique et des matériaux diélectriques LEMD du centre national de la recherche scientifique CNRS de Grenoble en France. La base de données consiste des caractéristiques courant-tension de la décharge couronne en configuration pointe-plan pour le SF₆ et le mélange SF₆-N₂. Les pressions utilisées varient entre 2 et 15 bars, les mesures sont effectuées pour la polarité négative et positive.

Nous avons scindé notre travail de recherche en 5 chapitres.

Le premier chapitre est consacré aux généralités sur le SF₆, ses avantages dans les systèmes d'isolation et ses inconvénients qui consistent à son rôle destructeur de l'environnement (gaz à effet de serre). Le remplacement du SF₆ devient une urgence et le mélange SF₆-N₂ est le gaz de substitution le plus probable pour le court terme.

Dans le deuxième chapitre, une étude rétrospective sur la théorie des décharges électriques dans les gaz avec une attention particulière aux phénomènes de conduction électrique.

Le troisième chapitre, est consacré aux résultats expérimentaux de mesures des caractéristiques courant-tension (I-V) pour la détermination des mobilités des porteurs de charge en utilisant la

méthode indirecte en polarité positive et négative dans les différents mélanges et sous haute pression.

Le chapitre Quatre traite les différents aspects de la technique des réseaux neurones pour la détermination des mobilités des porteurs de charges du SF_6 et du mélange $\text{SF}_6\text{-N}_2$.

Le chapitre 5 est consacré au travail réalisé. Application de réseaux de neurones pour la prédiction des mobilités dans les mélanges $\text{SF}_6\text{-N}_2$

Finalement la thèse se termine par la présentation des résultats de simulations, les discussions et la conclusion générale.

Chapitre 1

Généralités sur le SF₆

1.1 Introduction

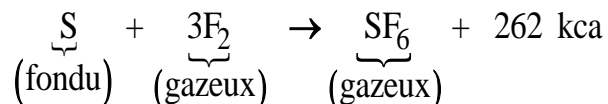
Les gaz constituent un des éléments de base de l'isolation à haute tension des matériels électriques. De tout temps, les bonnes qualités isolantes de l'air ont été utilisées aussi bien dans les applications de l'électricité statique que dans l'isolation des lignes aériennes. Dans ces dernières décennies, la nécessité de diminuer sensiblement les dimensions des installations électriques pour des questions d'encombrement et de coût a conduit à recourir à d'autres gaz isolants tels que l'azote, le CO₂, et surtout l'hexafluorure de soufre (SF₆). Parallèlement, malgré des difficultés pratiques importantes, on a utilisé les bonnes caractéristiques isolantes du haut vide ($10^{-5} - 10^{-7}$ mmHg) dans la technique de coupure (disjoncteurs, relais Reed).

Le comportement des isolants gazeux en haute tension doit être étudié pour deux raisons.

- éviter le développement d'arcs, ou si cela n'est pas possible, limiter ses effets (distances d'isolement, cornes et raquettes d'amorçage en parallèle avec des chaînes ou des fûts d'isolateurs, interrupteurs, disjoncteurs, fusibles, etc.);
- connaître la physique des amorçages d'arc dans les gaz afin de faciliter leur amorçage dans les applications particulières (éclateurs d'amorçage de générateurs de choc, éclateurs de protection, para surtensions, parafoudres, tubes fluorescents, soudure électrique, électroérosion,

1.2 Historique sur le SF₆

L'hexafluorure de soufre (SF₆) fut fabriqué pour la première fois en 1900 par H. Moissan et P. Lebeau [I-1] à partir du fluor et du soufre par la réaction suivante:



Les premières recherches d'envergure furent menées en 1939 par H.G. Pollak et F.S. Cooper [I.2] qui étudièrent le comportement électrique du SF₆ en champ divergent sous tension continue. Au cours des années suivantes les recherches portèrent principalement sur la stabilité chimique du gaz soumis à des décharges électriques. Ce n'est qu'aux années cinquante qu'on s'est intéressé au SF₆ comme un gaz isolant dans le domaine de l'électrotechnique de haute tension.

1.3 Propriétés du SF₆

1.3.1 Propriétés physico-chimiques du SF₆

La représentation de la molécule du SF₆ sur la figure 1, montre une structure dans laquelle l'atome de soufre occupe le centre d'un octaèdre régulier dont chaque sommet est occupé par un atome de fluor. C'est une configuration parfaitement stable, avec des liaisons saturées, qui confèrent à la molécule une grande inertie chimique. Les liaisons soufre-fluor sont fortement covalentes .

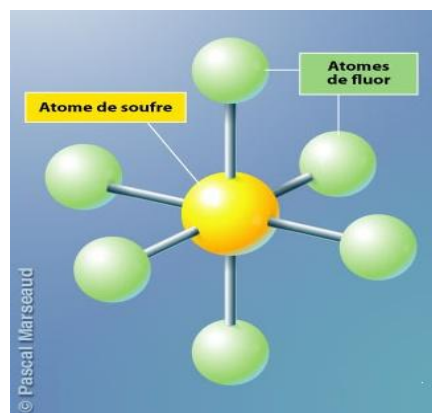
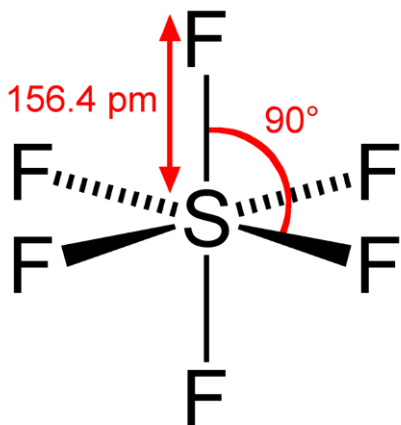


Figure 1.1: Structure moléculaire de l'hexafluorure de soufre (SF₆)

La structure moléculaire du SF₆ est à l'origine de l'excellente stabilité thermique, son électronégativité lui permet de neutraliser les électrons libres présents dans le milieu. Dans son état normal il est chimiquement neutre, non-toxique, et il est non inflammable. Le SF₆ a une capacité remarquable de s'auto-cicatriser après une rupture de l'intervalle, il est incolore, inodore et il est très compatible avec les matériaux utilisés dans l'industrie de l'électrotechnique. Sur le tableau I-1, on trouve les principales caractéristiques du SF₆.

Tableau 1-1 - Principales caractéristiques du SF₆

Masse molaire moléculaire	146,050418 g.mol ⁻¹
Masse volumique	6,602 kg.m ⁻³ à T=273,15 K et P=101,3 kPa
Point de sublimation	- 63,8°C à P=101,3 kPa
Point de fusion	- 50,5°C à P=101,3 kPa
Chaleur spécifique	C _p =0,159 kcal.kg ⁻¹ .K ⁻¹ à T=298,15K et P=101,3 kPa
Conductivité thermique à P=101,3 kPa	$\kappa=2,88.10^{-5}$ cal.cm ⁻¹ .s ⁻¹ .K ⁻¹ à T=273,15 K $\kappa=3,14.10^{-5}$ cal.cm ⁻¹ .s ⁻¹ .K ⁻¹ à T=288,15 K $\kappa=3,80.10^{-5}$ cal.cm ⁻¹ .s ⁻¹ .K ⁻¹ à T=323,15 K $\kappa=4,72.10^{-5}$ cal.cm ⁻¹ .s ⁻¹ .K ⁻¹ à T=373,15 K
Permittivité relative à P=94,4 kPa	ε _r =1,00191 à T=298,15 K

Décomposition du SF₆ dans les différents types de décharges électriques

1- DECOMPOSITION DANS UNE DECHARGE D'ARC:

Les principaux produits issus de la décomposition en présence d'une décharge d'arc sont: SOF₂; SF₄ et SF₂. [I.5]:

2- DECOMPOSITION DANS UNE DECHARGE ETINCELLE:

Les produits détectés par différents travaux [I.6 ; I.7] sont rassemblés dans le tableau 2:

Tableau I-2. Principaux produits issus de la décomposition du SF₆ sous décharge étincelle.

Produits	Concentration en % de volume
SOF ₂ (SF ₄)	0.5
SOF ₄	0.085
S ₂ F ₁₀	0.026
SO ₂ F ₂	0.006
SO ₂	0.002
HF	

3- Décomposition dans une décharge couronne :

Le SOF₂ a été fortement détecté ainsi que le SOF₄ et le SO₂F₂. Le S₂F₁₀ est un produit très toxique a été aussi détecté sous décharge couronne [I.8].

4- Décomposition thermique:

Le taux de décomposition du SF₆ augmente avec l'augmentation de la température et les composés qui peuvent apparaître sont: SOF₂, SO₂F₂, SO₂ (à 650°C).

5- Toxicité des produits de décomposition du SF₆:

Pur, le SF₆ est non toxique et inerte. Il peut cependant être asphyxiant lorsque sa concentration est élevée dans l'air. Le taux maximal conventionnellement admis est de 12 mg.m⁻³ [I.8]. Sur le tableau 3 sont présentés quelques produits de décomposition et leurs degrés de toxicité:

Tableau. I-3. Les principaux produits de décomposition [I.9].

Sous produits	Degré de toxicité	Quantité admissible (mg/m ³)
SOF ₂	Peu toxique	2.5
SOF ₂	Peu toxique	2.5
SF ₄	Moyennement toxique	0.1
SO ₂ F ₂	Moyennement toxique	5
SO ₂	Moyennement toxique	2
HF	Moyennement toxique	3
S ₂ F ₁₀	Très toxique	0.025

1.3.2 Propriétés électriques

Rigidité diélectrique

La « rigidité diélectrique » d'un gaz est la capacité de ce dernier à résister à une tension élevée. Ainsi, pour un gaz à une pression donnée, il existe une valeur de la tension, dite tension disruptive, au-dessous de laquelle le gaz reste isolant et au-dessus de laquelle il devient conducteur. Pour tester la rigidité diélectrique d'un gaz, on le place entre deux conducteurs séparés d'une petite distance (par exemple: deux sphères conductrices identiques peu écartées formant un spintermètre). On fixe la pression et on augmente progressivement la tension dans le dispositif. On constate que, pour une tension donnée, le gaz passe de l'état isolant à celui de conducteur et un courant circule entre les deux éléments conducteurs formant le système. Plus cette tension de claquage est élevée plus le gaz en question est un bon diélectrique.

Le SF₆ possède une très bonne tenue diélectrique. Lorsque le champ électrique est homogène et sous pression atmosphérique, le SF₆ a une tenue diélectrique d'environ trois fois meilleure que l'air (Figure. 1.2). Le SF₆ garde toujours sa bonne performance diélectrique en champ non uniforme ou en présence de diélectriques solides. En effet, dans les champs non uniformes, ce rapport atteint 5 fois à 0,2 MPa. Ainsi, les distances d'isolement, ainsi que la taille des équipements isolés au SF₆, s'en trouvent considérablement réduites, par rapport à leurs homologues isolés à l'air. En outre, sa chaleur spécifique est 3,7 fois supérieure à celle de l'air, ce qui limite significativement l'échauffement des équipements [I.3].

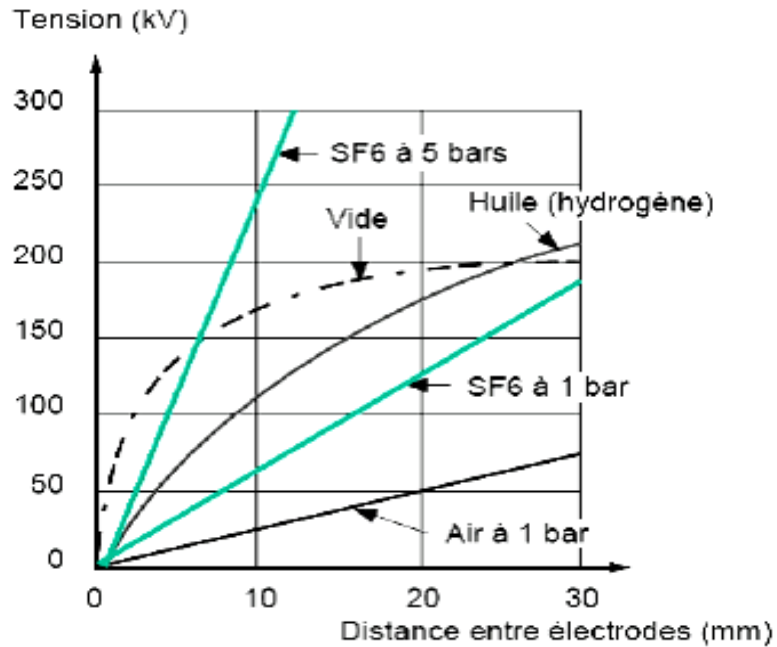


Figure 1.2: Tenue diélectrique du SF₆ comparé à différents matériaux en champ uniforme [I.3]

Le SF₆ a un point d'ébullition raisonnablement bas, qui est d'environ -64°C [I.4]. Il est donc possible d'utiliser le SF₆ comprimé à quelques bars sans risque de liquéfaction du gaz même en hiver. Par exemple, le SF₆ comprimé à 3 bars conserve toujours son état gazeux même si la température chute à -40°C

1.3.3 Le SF₆ et l'atmosphère

L'Hexafluorure de soufre est depuis longtemps largement utilisé comme gaz d'isolation dans les appareillages moyenne et haute tension (disjoncteurs, postes sous enveloppe métallique, câble à isolation gazeuse...). Ce gaz possède d'excellentes qualités techniques telles qu'une bonne tenue diélectrique, un point d'ébullition bas ou encore une bonne stabilité chimique. Toutefois, Figure 1.3, montre une tendance continue d'augmenter avec un taux de croissance encore plus en plus, il a été mis en évidence que le SF₆ est un gaz à effet de serre notable. Son coefficient GWP100 (Global Warming Potential sur une durée de 100 ans) est de 24000. Ce qui veut dire qu'un kilogramme de SF₆ au même impact sur le réchauffement climatique que celui de 24000 kg de CO₂. Sa durée de vie dans l'atmosphère est très longue, estimée entre 800 et 3200 ans [I.10].

La contribution du SF₆ au réchauffement terrestre pose le problème de la réduction de son utilisation à court terme, et de son remplacement à moyen terme. figure1.3

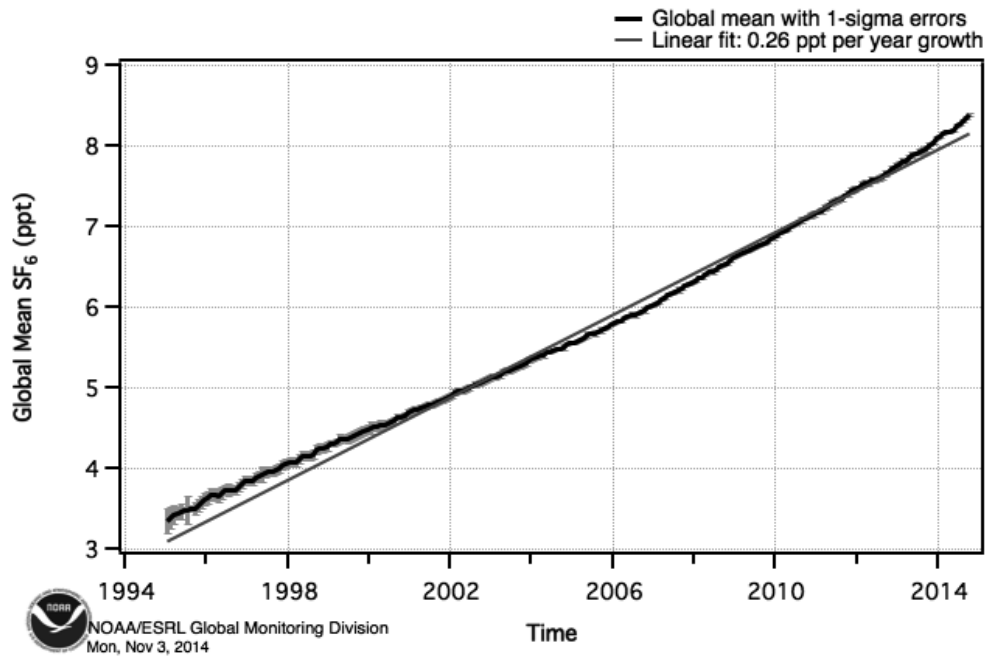


Figure 1.3: Le développement de la teneur atmosphérique moyenne mondiale de SF₆.

La recommandation CEI 62271 *Appareillage à haute tension*, précise les conditions d'étanchéité requises pour les installations à gaz. Le taux de fuite F_p est défini comme la masse de gaz qui s'échappe de l'installation, en pourcent de la masse totale du gaz, par année [% p.a.]. La recommandation de la CEI distingue les « systèmes fermés » et les « systèmes étanches » (soudés). Les directives actuelles spécifient, pour les installations de grandes dimensions isolées au SF₆ :

$$F_p < 0,5\% \text{ p.a.} \quad (\text{Systèmes fermés})$$

$$F_p < 0,1\% \text{ p.a.} \quad (\text{Systèmes étanches})$$

Ces taux ont été définis de manière à limiter autant que possible la quantité de gaz rejeté dans l'atmosphère durant toute la durée de vie de l'installation, qui est de l'ordre de 20 à 40 ans. Les procédures de contrôle du taux de fuite sont spécifiées par la norme CEI 60068-2-17, [commission électrotechnique internationale, Essais fondamentaux climatiques et de robustesse mécanique] – Essai Q : étanchéité.]

1.4 Utilisation des mélanges de SF₆

Le remplacement du SF₆ a fait l'objet de nombreuses recherches depuis les années 1970. Il a été montré qu'il est très difficile, voire impossible, de trouver un gaz « vert » capable de remplacer universellement le SF₆. Les solutions pour son remplacement dépendent fortement de l'application envisagée. Un autre axe de réflexion consiste à réduire la quantité de ce gaz en utilisant des mélanges SF₆ / autre gaz compatible avec l'environnement, avec une teneur en SF₆ la plus faible possible. Ceci consiste à ajouter un faible taux de SF₆ dans un autre gaz comme N₂, CO₂, air ou encore hélium, argon [ou hydrogène [I.11-I.13]]. Le mélange SF₆-H₂ est inutilisable à cause du danger d'explosion en cas de décharge électrique [I.14]. Les mélanges SF₆-Ar et SF₆-He ont un champ critique réduit nettement inférieur à celui du mélange SF₆, et ils n'ont pas d'intérêt industriel sauf éventuellement le mélange SF₆-He pour la coupure d'arc. En conclusion pour l'isolation électrique, seuls les mélanges SF₆ avec N₂, CO₂ ou air sont intéressants, car un effet de synergie est observé avec ces trois mélanges (Fig. I.3) pour des concentrations faibles en SF₆ [I.15].

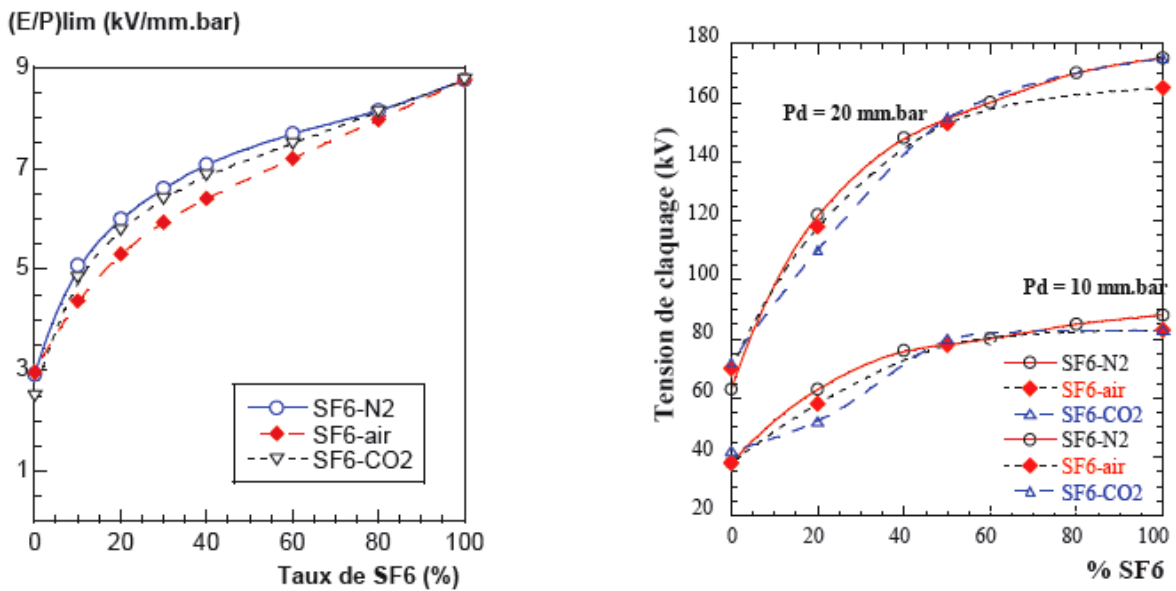


Figure 1.4: Evolution du champ réduit critique $(E/P)_{lim}$ (courbes construites de [I.16]) et tension de claquage en champ uniforme (résultats expérimentaux) de mélanges de SF₆ avec N₂, CO₂ et air en fonction du taux de SF₆ (courbes construites de [I.16 - I.17])

L'effet de synergie des mélanges de SF₆ avec N₂, air et CO₂ a été montré aussi bien par la mesure du champ critique que par la mesure de la tension de claquage (Figure I.3). L'accroissement de la performance diélectrique est très important aux faibles taux de SF₆. Par exemple, un mélange 20% SF₆ – 80% N₂ a une tension de claquage deux fois meilleure que celle du N₂ pur. Cet effet de synergie est particulièrement favorable car il permet de réduire la quantité de SF₆ utilisée en conservant une tenue diélectrique nettement améliorée par rapport à N₂, air ou CO₂ pur. Du point de vue économique, ces mélanges sont également intéressants car N₂, air ou CO₂ sont abondants dans l'atmosphère. A ce jour, les mélanges SF₆-N₂ à faible taux de SF₆ (< 20%) sont les plus étudiés. Ces mélanges ont été utilisés avec succès dans l'industrie électrique. Par exemple, dans certains disjoncteurs où un fonctionnement à très basse température est exigé [I.18]. Plus récemment, il a été utilisé pour les lignes de transport à isolation gazeuse (LIG) de seconde génération [I.19], qui sont commercialisées depuis 2001 (LIG 220 kV à Palpexpo à Genève, Suisse) [I.21].

1.5 Propriétés de l'azote (N₂)

1.5.1 Propriétés physico-chimiques

L'azote est le constituant majoritaire de l'atmosphère terrestre, environ 78 % en volume. Il est incolore, inodore, inflammable, non corrosif et non toxique. Toutefois, lorsqu'il est en très grande quantité, il peut être asphyxiant (par déplacement de l'oxygène).

Les propriétés physico-chimiques de l'azote sont résumées dans le tableau I-4 ci-dessous [I.13].

Tableau I-4: Principales propriétés du N₂

Paramètres	Valeur	Conditions
Masse molaire moléculaire	28,0134 g.mol ⁻¹	/
Masse volumique de gaz	1,25053 kg.m ⁻³ [Ia-1]	T.P.N.
Point de fusion	63 K	P=101,3 kPa
Point d'ébullition	77,347 K	P=101,3 kPa
Energie de première ionisation	E _i = 15,58 eV	
Constante diélectrique	1	à T = 25°C et p = 1bar
Facteur de perte diélectrique tan (δ)	= 0 (inférieure à 2.10 ⁻⁷)	

Masse spécifique	1,18 kg/m ³	à p= 1 bar et T = 20°C
Dimension moléculaire	3 à 4 Å	
Point triple	T = -210 °C; p = 0,12 bars; L _f = 6,15 kcal/kg	
Point critique	T = -146,4 °C; p = 33,99 bars ; ρ = 314 kg/m ³	
Facteur de compressibilité Z(pV/RT)	0,9984	à T = 27 °C et p = 10 bars
	0,9965	à T = 7 °C et p = 10 bars
	0,9991	à T = 27 °C et p = 5 bars
Conductivité thermique (mW/cm.°C)	0,258	à T = 27°C et p = 1 bars
	0,244	à T = 7°C et p = 1 bars
	0,262	à T = 27°C et p = 10 bars

Sous forme gazeuse, l'azote ne se liquéfie qu'à de très basses températures : - 196°C pour une pression de 1 bar. Cette propriété lui confère un avantage indéniable sur le SF₆ pour une utilisation dans les régions de climat très froid. La conductivité thermique de l'azote est plus élevée que celle du SF₆. Ce gaz isolant est effectivement adapté pour une éventuelle utilisation sur une très grande échelle de température. La rigidité diélectrique de l'azote est 2 à 3 fois plus faible que celle du SF₆.

Afin d'égaliser les performances diélectriques du SF₆ on peut miser sur l'augmentation de la pression d'utilisation de l'azote, on a envisagé a priori une pression d'azote environ trois fois supérieure à la pression usuelle de SF₆ pour atteindre la même tenue diélectrique.

La contribution majeure de l'azote dans le mélange SF₆-N₂ est de ralentir les collisions électroniques en réduisant l'énergie des électrons. Le processus vibratoire intense tend à renvoyer les électrons dans la zone où la probabilité d'attachement du SF₆ est élevée. L'effet de synergie de l'azote a été mis en évidence et l'abondance de l'azote dans l'air lui donne un atout supplémentaire du point de vue coût.

Les avantages d'un tel mélange résident dans:

La réduction du coût du système.

L'utilisation d'une faible concentration du SF₆ dans le mélange fournit une solution au problème de la contribution de ce dernier au réchauffement de l'atmosphère.

Elargissement de l'utilisation de l'isolant aux zones très froides et à des pressions relativement élevées sans pour autant provoquer la liquéfaction du diélectrique. Ceci peut être vérifié par les courbes de pression de vapeur représenté sur la figure iv, pour le SF₆, N₂ et le mélange SF₆-N₂.

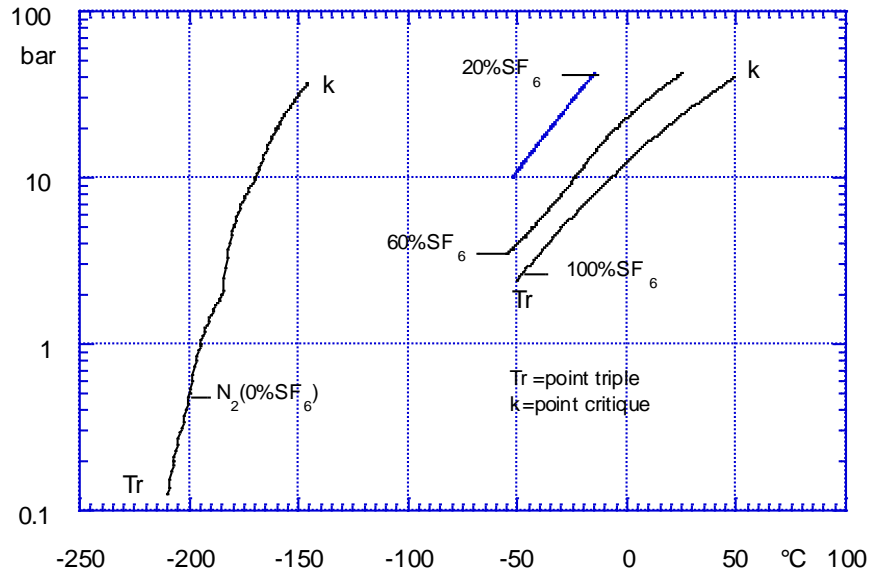


Figure 1.5: Courbes de pression de vapeur pour le SF₆, N₂ et SF₆-N₂ [I.21].

On peut constater que le SF₆ se liquéfie à des températures nettement supérieures à celles de l'azote. Avec le mélange SF₆-N₂ ou le gaz idéal est l'azote et le SF₆ est considéré comme gaz condensable, la liquéfaction du mélange dépend de la pression.

1.6 Facteurs influençant la tenue diélectrique de l'azote

A l'instar du SF₆, le comportement de l'azote est assujéti aux facteurs qui modifient le champ ou les phénomènes photoniques dans l'intervalle. Toutefois, la croissance du coefficient primaire d'ionisation en fonction du champ est relativement moins rapide dans l'azote que dans le SF₆. Les effets des facteurs agissant sur le champ sont comparativement moins prononcés dans un intervalle d'azote que dans celui de SF₆. C'est entre autres le cas des particules métalliques fixes et libres. Les matériaux des électrodes, en intervenant dans la décharge par le biais des mécanismes secondaires d'émissions suite aux bombardements ioniques ou photoniques, auront une influence non négligeable à cause du rôle prépondérant de la photo ionisation dans l'azote.

Quant aux impuretés gazeuses, leur impact sera lié à leur comportement vis à vis des phénomènes photoniques. La tenue de l'intervalle sera réduite par les molécules des impuretés si leur énergie d'ionisation et/ou leur énergie d'émission de photons sont inférieures à celles de la molécule d'azote (neutre ou excitée).

1.7 Propriétés du mélange SF₆-N₂

L'utilisation du SF₆ dans des quantités plus faibles en le mélangeant avec un autre composant a été envisagée afin de limiter ainsi ses rejets dans l'atmosphère qui contribuent à l'effet de serre. L'idéal serait de le remplacer par un gaz équivalent au niveau diélectrique mais sans conséquences sur l'environnement. Depuis quelques années, des diélectriques gazeux meilleurs ou équivalents ont été étudiés (perfluorocarbones et leurs mélanges par exemple) mais présentent d'autres inconvénients sérieux. Le meilleur compromis à ce jour a été de mélanger le SF₆ avec un gaz tel que le diazote car même avec de très faibles quantités d'hexafluorure de soufre, le mélange SF₆-N₂ présente une bonne tenue diélectrique. Le diazote transporte les électrons énergétiques dans les niveaux d'énergie les plus bas (on peut noter que le CO et le CO₂ jouent le même rôle que N₂ [I.22 ; I.23]. où le SF₆ les capture avec une très forte probabilité. La création d'autres électrons dans le gaz est ainsi évitée de même que le phénomène d'ionisation et par la suite le claquage.

A long terme, l'élimination totale du SF₆ est à prévoir même si cette solution n'est pas encore envisageable aujourd'hui pour des raisons techniques et économiques.

Les mélanges SF₆-N₂ à faible concentration en SF₆ (5 à 15 %) sont destinés au domaine des lignes de transmission comme gaz isolant sous des pressions comprises entre 500 et 900 kPa [I.24]. Leur utilisation va évidemment impliquer une augmentation du coût puisque le recyclage requiert une adaptation nécessaire des installations mais ceci est déjà envisagé par plusieurs pays [I.25 ; I.26].

Dans l'industrie, les tendances actuelles sont l'utilisation du SF₆ pur pour les disjoncteurs mais de plus en plus de soins et d'attention sont apportés à la réduction des fuites ainsi qu'à la récupération et au recyclage du gaz, alors que pour l'isolation on s'oriente vers les mélanges gazeux comme SF₆-N₂ (≅ 10%-90%) surtout dans des conditions de température extrêmes où le SF₆ pur n'est plus adapté.

1.7.1 Produits de décomposition gazeux formes dans les mélanges SF₆-N₂ soumis a des décharges couronne de polarité négative

Les travaux effectués par Castonguay et Dionne [I.27] sur des mélanges SF₆-N₂ (50%-50%) soumis à des décharges couronne de polarité négative ont montré que le diazote induit la production, dans des quantités significatives, de quelques sous-produits nouveaux comme SF₅N₂F, (SF₅)₂NF, SF₃NF₂, SF₅-N₂-SF₅, SO₂F-NF-SOF, NSF₃. A cette liste viennent s'ajouter les sous-produits issus seulement du diazote comme les oxydes d'azote (NO, NO₂, N₂O) et des fluorures d'azote (NF₃, N₂F₄) formés à partir de fragments issus des deux composants du mélange. D'autres sous-produits ont également été détectés par Castonguay et Dionne [I.28] mais en si petite quantité qu'ils présentent peu d'intérêt. Le fait que le nombre de sous-produits de décomposition soit bien plus important dans les mélanges SF₆-N₂ que dans le SF₆ pourrait accroître les problèmes de toxicité.

Christophorou et Van Brunt ont indiqué que dans les mélanges SF₆-N₂ (même à forte teneur en N₂) soumis à des décharges couronne, les sous-produits principaux restent ceux formés dans le cas du SF₆ pur. Ils ont, par ailleurs, constaté que le taux de dissociation du SF₆ dans ces mélanges, bien qu'il n'augmente pas considérablement quand la teneur du mélange en diazote passe de 0 à 40%, augmente d'une manière dramatique dès que la teneur en diazote dépasse 80%. Cela provient d'une diminution du taux de recombinaison des fragments de dissociation du SF₆.

1.7.2 Décomposition sous l'effet de claquages de mélanges SF₆-N₂

La décomposition sous l'effet de claquages de mélanges SF₆-N₂ conduit comme sous décharges couronnes, aux mêmes sous. Produits principaux que ceux formés dans le cas du SF₆ pur et ceci même pour des concentrations de diazote égales ou supérieures à 90% [I.28].

1.8 Conclusion

Le remplacement du SF₆ par un gaz équivalent est pratiquement impossible dans le cours terme, le meilleur compromis est de mélanger le SF₆ avec un gaz tel que le diazote avec de très faibles

quantités d'hexafluorure de soufre. Le mélange SF₆-N₂ présente une bonne tenue diélectrique au niveau diélectrique mais sans conséquences sur l'environnement.

Bibliographie

CHAPITRE I

- [I.1] H. Moisson et P. Lebeau, *Comp. Rev.* 130, pp. 865, 1900.
- [I.2] H. G. Pollak et F. S. Cooper, *physical revue* 56, pp. 170-175, 1939.
- [I.3] S. Théoleyre, *Techniques de coupure en moyenne tension*. Techniques de l'ingénieur D4705, 1999.
- [I.4] N. Marshall, *Encyclopedie des gaz Gas encyclopaedia translated by Nissim Marshall*. Amsterdam: Elsevier Scientific Pub., 1976.
- [I.5] F.Y. Chu. "SF₆ Decomposition in Gas-Insulated Equipment", *IEEE. Trans. on Elect. Insul.* Vol. EI. 21, N°5, 1986.
- [I.6] G. D. Griffin et al "Concerning biological effect of sparked decomposed SF₆", *I.E.E Proceeding*, Vol. 137. N°4, pp. 221-227, 1990.
- [I.7] I. Sauers et al. "Production of S₂F₁₀ in sparked SF₆" *J. Phys D: Appl. Phys.* Vol. 21, pp1230-1238, 1988.
- [I.8] D. R. James et al. "Investigation of S₂F₁₀ production and mitigation in compressed SF₆-insulated power systems", *IEEE Electrical insulation magazine*. Vol. 9. N°3, pp 29-40, 1993.
- [I.9] P. Segur "gaz isolants", *Technique de l'ingénieur, Traité de Génie Electrique*, Vol. D. 2530-2531, 1990.
- [I.10] L. Christophorou, J. Olthoff, and R. Van Brunt, "Sulfur hexafluoride and the electric power industry," *IEEE Electrical Insulation Magazine*, vol. 13, no. 5, pp. 20-24, 1997.
- [I.11] Y. Qiu and X. Weng, "Prediction of dielectric strength of SF₆/He gas mixtures," *Journal of Physics D: Applied Physics*, vol. 23, no. 3, pp. 363-364, 1990.
- [I.12] W. Khechen and J. Laghari, "Breakdown studies of SF₆/argon gas mixtures," *IEEE Transactions on Electrical Insulation*, vol. 24, no. 6, pp. 1141-1146, Dec. 1989.
- [I.13] J. L. Moruzzi and J. D. Craggs, "Ionization and attachment and breakdown measurements in mixtures of SF₆ with helium, argon and hydrogen," in *Proceedings of the Twelfth International Conference on Phenomena in Ionized Gases*, 1975..
- [I.14] N. Malik and A. Qureshi, "A Review of Electrical Breakdown in Mixtures of SF₆ and Other Gases," *Electrical Insulation, IEEE Transactions on*, vol. 14, no. 1, pp. 1-13, 1979.

- [I.15] Y. Qiu and E. Kuffel, "Comparison of SF₆/N₂ and SF₆/CO₂ gas mixtures as alternatives to SF₆ gas," *Dielectrics and Electrical Insulation, IEEE Transactions on*, vol. 6, no. 6, pp. 892-895, 1999.
- [I.16] N. Malik and A. Qureshi, "Breakdown Gradients in SF₆-N₂, SF₆-Air and SF₆-CO₂ Mixtures," *Electrical Insulation, IEEE Transactions on*, vol. 15, no. 5, pp. 413-418, 1980.
- [I.17] M. Siddagangappa, C. Lakshminarasimha, and M. Naidu, "Prediction of Breakdown Voltages of Binary Gas Mixtures in Uniform Electric Fields," *Electrical Insulation, IEEE Transactions on*, vol. 19, no. 4, pp. 359-361, 1984.
- [I.18] A. Buttka, J. Pfeiffer, and B. Sojka, "Siemens SF₆/N₂ Circuit-Breaker for Service at Low Temperatures," *Siemens Power Engineering, Special Issue "HV Technology"*, vol. 6, pp. 28-37, 1984.
- [I.19] H. Koch and T. Hillers, "Second generation gas-insulated line," *Power Engineering Journal*, vol. 16, no. 3, pp. 111-116, 2002.
- [I.20] F. Renaud, "220 kV gas-insulated transmission line-Palexpo Geneva Switzerland," in *2003 IEEE Power Engineering Society General Meeting (IEEE Cat. No.03CH37491)*, pp. 2476-2479, 2003.
- [I.21] J. C. Devins "Replacement gases for SF₆", *IEEE. Trans. on Elec. Insul. Vol. E. I15, N°2*, pp 81-85, April 1980.
- [I.22] L.G. Christophorou, J.K. Olthoff et D.S. Green « Gases for Electrical Insulation and Arc Interruption: Possible Present and Future Alternatives to Pure SF₆. ». National Institute of Standards and Technology (N.I.S.T). Technical Note **1425**, (1997).
- [I.23] L.G. Christophorou « Insulating gases ». *Nuclear Instruments and Methods in Physics Research, A268*, pp. 424-433, (1988).
- [I.24] H. G. Pollak et F. S. Cooper, *physical revue* 56, pp. 170-175, 1939.
- [I.25] O. Yamamoto, T. Takuma, A. Kawamura, K. Hashimoto et M. Kinouchi « SF₆ gas recovery from SF₆/N₂ mixtures using polymers membrane ». *Gaseous Dielectrics IX*, eds : L.G. Christophorou and J.K. Olthoff, Kluwer Academic / Plenum Publishers, NY, pp. 555-560, (2001).
- [I.26] M.Pittroff et F. Vondenhof « Separation of SF₆/N₂ mixtures ». *Gaseous Dielectrics IX*, eds : L.G. Christophorou and J.K. Olthoff, Kluwer Academic / Plenum Publishers, NY, pp. 561-566, (2001).
- [I.27] J.Castonguay et I.Dionne « S2F10 and other heavy gaseous decomposition byproducts formed in SF₆ and SF₆-gas mixtures exposed to electrical discharges ». *7th International Symposium on Gaseous Dielectrics*, Knoxville, 24-28 Avril, (1994).

Chapitre 2

Etude de la décharge couronne

2.1 Introduction

Les phénomènes de décharge électrique, étudiés pourtant depuis près de deux siècles, font toujours l'objet de recherche active auprès des scientifiques. On pourrait s'étonner de cet état de fait mais on comprendrait vite que cela n'est dû qu'à la très grande complexité du phénomène et des mécanismes qui le constituent. Cependant de grandes avancées ont été faites ces dernières décennies dans l'amélioration des moyens de diagnostic des premiers instants de la décharge et donc dans l'étude des mécanismes d'initiation. Cela a permis d'accroître les connaissances sur le phénomène et d'avancer dans les applications qui ont d'ailleurs connu de bien plus larges expansions.

Dans ce chapitre, nous élaborerons quelques grandes théories sur les mécanismes d'initiation d'une décharge électrique dans les gaz, en élaborant sommairement les théories allant des électrons primaires en passant par les avalanches jusqu'à la rupture de l'intervalle.

2.2 Processus de formation des porteurs de charge dans le SF₆

2.2.1 Formation des électrons primaires

Il est tout à fait bien connu qu'aucune décharge électrique ne peut s'initier sans la présence d'au moins un électron dans le volume de gaz couvrant l'espace inter-électrodes. Ces électrons, présents au tout début de la décharge sont communément appelés électrons germes ou encore électrons primaires. Ils permettent d'initier la première avalanche électronique sans laquelle aucun type de décharge ne peut prendre forme encore moins se maintenir. Ils peuvent être produits par

radiation cosmique, par les sources naturelles d'ionisation, par les charges résiduelles issues des décharges précédentes ou encore par détachement des ions négatifs par effet de champ. De ces sources de production d'électrons, il peut y avoir jusqu'à 20 paires d'ion-électron par mm^3 par seconde. Dans la région d'accentuation du champ électrique, ces électrons primaires seront accélérés et pourront gagner de leur parcours une certaine énergie d'ionisation de quelques eV avant leur première collision. L'énergie nécessaire pour ioniser une molécule d'air (20% de O_2 , 80% de N_2) est d'environ une quinzaine d'eV (12.5 pour l'oxygène et 15.57 pour l'azote) [III.1], au moins un des électrons primaires accélérés par le champ pourrait donc gagner l'énergie suffisante pour provoquer la première ionisation.

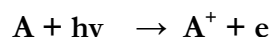
2.2.2 Ionisation, recombinaison et attachement électronique

L'existence d'une décharge électrique résulte du passage d'un courant dans le diélectrique gazeux et donc de la possibilité de créer des particules chargées, essentiellement des électrons, à partir des atomes ou molécules neutres du gaz. Les principaux mécanismes régissant la génération des porteurs de charge dans les gaz sont:

Les mécanismes d'ionisation, tendant à augmenter leur densité et les mécanismes de recombinaison et d'attachement, tendant à les réduire.

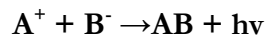
Le phénomène d'ionisation est un processus tout à fait aléatoire mais il peut être formellement lié au *libre parcours moyen pour l'ionisation*. Le processus est largement décrit dans la littérature scientifique et on peut noter que la nature du gaz et sa densité sont les paramètres les plus influents.

Un électron peut ioniser une particule neutre par collision à condition qu'il ait une énergie supérieure à celle d'ionisation de la particule. Cependant à température élevée, surtout avec une forte pression, le gaz de remplissage peut s'ioniser lui-même et fournir ainsi d'autres électrons libres. Un photon suffisamment énergétique, généralement libéré par un atome métastable retrouvant son état d'équilibre peut également contribuer à l'ionisation des particules neutres selon le principe suivant.



D'autre part, l'ionisation peut être également produite par les ions positifs, mais ce processus exige que la particule incidente ait un haut niveau énergétique, de l'ordre de 100 ou 200 eV. Le phénomène n'intéresse que certains types de décharges à basse pression, dans lesquelles les ions accélérés par le champ électrique, peuvent acquérir une énergie importante entre deux collisions. Dans tous les cas, lors de la collision, l'excédent d'énergie de l'électron, de la particule, du photon ou de l'ion est fourni à l'électron nouvellement créé sous forme d'énergie cinétique.

La *recombinaison* a été beaucoup moins étudiée que son processus inverse, l'ionisation. Cependant elle reste d'une importance fondamentale pour la bonne compréhension des mécanismes de développement de la décharge et tout particulièrement pour ceux de sa phase d'extinction. La recombinaison est la neutralisation d'un ion positif par capture d'un électron ou d'ion négatif suivant l'équation ci-dessous.

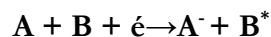
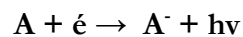


Le taux des recombinaisons ion - ion est généralement beaucoup plus élevé que celui des recombinaisons électron - ion, ce qui est dû à la faible mobilité des ions positifs.

Les phénomènes de recombinaison sont en général négligés dans les calculs des paramètres des décharges électriques mais sous certaines conditions fixées évidemment par les caractéristiques du gaz et la pression, ils ne peuvent plus l'être. Il en est de même de l'attachement électronique qui est le phénomène se produisant quand un électron rentre en collision avec un atome ou une molécule de gaz neutre pour former un ion négatif stable. L'état de stabilité n'est pas toujours atteint car l'ion négatif formé peut rentrer dans un état intermédiaire avant de se dissocier.

Par exemple O^- , O_2^- , NO_2^- , NO_3^- , OH^- , H^- et les ions halogènes négatifs sont formés de façon stable mais non N^- , N_2^- ou encore les ions négatifs des gaz rares .

L'*attachement* électronique est d'autant plus important que la molécule présente une plus grande affinité électronique. Le processus d'attachement des électrons est le suivant :



L'excédent d'énergie fourni par l'électron incident (c'est-à-dire la différence entre son énergie cinétique et l'affinité électronique de la molécule) peut se retrouver sous forme d'émission

radiative dans le cas d'une collision entre deux corps ou sous forme de source d'énergie cinétique pour une troisième particule dans le cas d'une collision à trois corps.

Dans le seul contexte des mécanismes d'initiation de la décharge, la recombinaison et l'attachement électronique peuvent jouer un rôle très important. Ils peuvent d'une part retarder l'initiation de la décharge car favorisant la formation d'états intermédiaires et d'autre part l'inhiber tout simplement. Il est certain que sous n'importe quelles conditions, une multiplication électronique ne peut s'initier que dans une région où les phénomènes d'ionisation sont plus importants que ceux de recombinaison et d'attachement. Dans l'air, cette condition est fournie par un champ réduit égal à: $E / P = 34 \text{V.cm}^{-1}.\text{torr}^{-1}$, soit 26kV.cm^{-1} à la pression atmosphérique [II.2].

2.3 Claquage diélectrique dans un gaz

Un gaz est théoriquement un isolant parfait puisqu'il ne contient que des atomes ou des molécules neutres. Pour que l'initiation de la décharge disruptive se produise, il faut qu'au moins une particule chargée dite particule germe, généralement un électron, soit présent dans ce gaz. Dès lors que cet électron est créé, il est accéléré par le champ électrique appliqué et entre en collision avec les molécules neutres du gaz. Quand le champ est suffisant, cette charge, entre deux chocs, prend assez d'énergie pour ioniser une molécule et créer d'autres charges libres qui à leurs tours peuvent ioniser d'autres molécules. C'est l'effet d'avalanche électronique. Plusieurs avalanches peuvent ainsi être créées dans le gaz. Ces dernières peuvent conduire au claquage électrique suivant deux mécanismes : le mécanisme de Townsend ou le mécanisme du canal (Streamer en anglo-saxon) en fonction de plusieurs paramètres tels que la distribution du champ électrique entre les électrodes, l'intensité de ce champ, la forme de la tension appliquée, le taux d'humidité, ... etc.

2.3.1 L'avalanche électronique

Considérons un système de deux électrodes planes et parallèles plongées dans un gaz isolant et entre lesquelles il y a un électron libre. Lorsqu'on applique une tension, l'électron est de plus en plus accéléré sous l'effet du champ électrique. Il parvient ainsi à gagner une énergie suffisante pour ioniser par collisions des particules neutres, donnant naissance à des ions positifs et à de nouveaux électrons. Ce phénomène est caractérisé par un coefficient a appelé coefficient d'ionisation (ou premier coefficient de Townsend).

Chaque électron nouvellement créé contribue à son tour, par ionisation par collisions au cours de son déplacement vers l'anode (électrode positive), à la création d'une avalanche électronique.

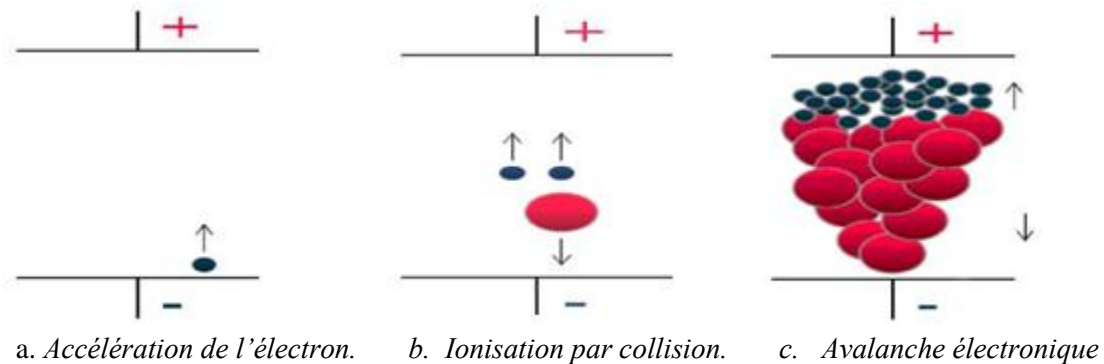


Figure 2.1: Processus d'avalanche électronique [II.3].

2.3.2 Décharge de Townsend

Nous On considère un gaz contenu entre deux électrodes, il peut y contenir, en l'absence de tout champ appliqué, un certain nombre de particules chargées. Elles sont produites soit par des mécanismes artificiels (par exemple en bombardant la surface de la cathode avec une source de rayonnement ultraviolet pour extraire des électrodes), soit par des mécanismes naturels (par exemple la radioactivité ou les rayons cosmiques). Quel que soit leur origine, ces charges primaires, en appliquant à l'espace interelectrodes une tension U donc un champ électrique E , vont entraîner la création, dans le circuit extérieur, d'un courant de très faible intensité, qui va croître avec la tension U .

Nous considérons un système d'électrodes planes, créant un champ homogène E , baignant dans un gaz à basse pression, figure (2.2). Un électron germe, accéléré par le champ électrique E appliqué, acquiert une énergie cinétique. A partir d'une certaine valeur de E , cette énergie suffira à ioniser, par collision, une particule neutre du gaz. Les nouveaux électrons créés vont participer,

à leur tour, au mécanisme d'ionisation des molécules neutres, produisant ainsi une multiplication électronique.

Si les électrons primaires sont produits par extraction à partir de la cathode, avec un taux de production N_0 , et si $N(x)$ est le nombre d'électrons produits par avalanche à la distance x de la cathode, le nombre dN de nouveaux électrons produit entre x et $(x + dx)$ sera donné par :

$$dN = N \cdot \alpha \cdot dx$$

α est appelé le premier coefficient de Townsend, il représente le nombre d'ions créés par un électron parcourant une distance d'un centimètre. Le coefficient α dépend de la nature du gaz, du champ local E et de la pression P du gaz, d'où :

$$N = N_0 \cdot e^{\alpha x}$$

Si d est la distance entre les électrodes, le courant mesuré à l'anode ou dans le circuit extérieur sera donné par:

$$I = I_0 \cdot e^{\alpha d} \quad (2.1)$$

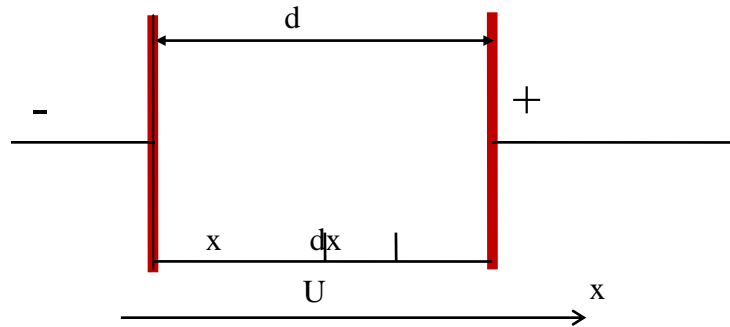


Figure 2.2: Configuration plan-plan

Si l'on fait croître la distance d à E/P constant, la croissance de $\ln(I/I_0)$ devient plus rapide qu'un accroissement linéaire à partir d'une certaine distance d . Ce comportement s'explique par l'action, sur la surface cathodique, des ions positifs laissés par la première avalanche. Ces ions dérivent vers la cathode, sous l'effet du champ électrique, et peuvent ainsi acquérir une énergie cinétique suffisante pour arracher des électrons de la surface cathodique. On définit alors un

autre coefficient γ , appelé deuxième coefficient de Townsend. Il représente le nombre moyen d'électrons arrachés à la cathode par les ions positifs incidents.

Les N_0 électrons primaires émis par la cathode donnent en retour $N_0 [e^{\alpha d} - 1]$ ions positifs. Ces derniers vont produire, par bombardement de la cathode, $\gamma N_0 [e^{\alpha d} - 1]$ électrons secondaires ; c'est à dire de nouveaux électrons qui se multiplieront comme les électrons initiaux. Le nombre d'électrons atteignant l'anode est donné par l'équation suivante:

$$N = N_0 \frac{e^{\alpha d}}{1 - \gamma \cdot (e^{\alpha d} - 1)} \quad (2.2)$$

Le critère de claquage de Townsend se traduit par l'existence d'une distance critique d pour laquelle le courant I croît vers l'infini. Ce critère s'écrit :

$$1 - \gamma \cdot (e^{\alpha d} - 1) = 0$$

Comme $e^{\alpha d} \gg 1$, nous pouvons écrire:

$$e^{\alpha d} = \frac{1}{\gamma}$$

Dans les zones à faible champ, l'attachement électronique ne peut plus être négligé. Ce phénomène est caractérisé par le coefficient d'attachement (η), qui s'oppose au processus d'ionisation (α). Il représente le nombre d'électrons captés, par les molécules du gaz, par unité de longueur. On définit alors le coefficient effectif d'ionisation $\bar{\alpha}$ comme suit :

$$\bar{\alpha} = \alpha - \eta$$

La théorie de Townsend fait intervenir tout l'intervalle interelectrode. Elle prévoit, à des pressions élevées, des temps de transition des charges entre les électrodes beaucoup plus grand que le temps de la formation de la décharge. Ceci permet d'écarter le processus secondaire « γ », car les ions positifs sont freinés et leur énergie cinétique n'est plus suffisante pour extraire des électrons supplémentaires de la surface cathodique. Les travaux de Fletcher [II4] confirment le non validité de la théorie de Townsend pour les pressions élevées du gaz et en champ uniforme.

Cependant, cette théorie conserve sa validité pour des valeurs du produit $p.d$ inférieures à 200 Torr.cm [II.5].

Le non validité de la théorie de Townsend, pour les pressions élevées et les grandes distances interelectrodes, a conduit à la recherche d'une autre théorie améliorée, celle de la formation du streamer.

2.3.3 Loi de Paschen

La tension d'alimentation minimale pour déclencher la décharge est appelée tension de claquage. Il existe une relation qui lie la valeur de cette tension de claquage à la pression du gaz et de la distance interélectrode : la loi de Paschen. Cette loi est habituellement représentée par la courbe caractéristique $U_c = f(p.d)$ où p est la pression, d la distance interélectrode, et dépend de la nature du gaz introduit.

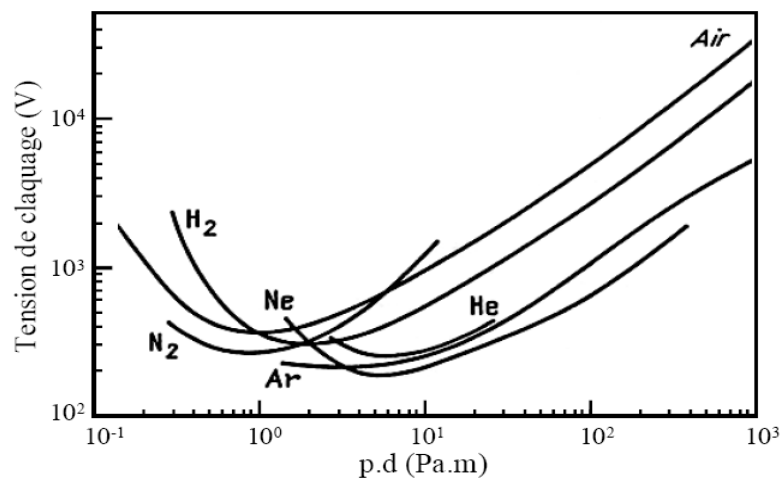


Figure 2.3: Schéma de la relation entre deux images Tension de claquage pour différents gaz en fonction du produit $p.d$ [II.6]

Nous pouvons constater que pour un gaz à pression atmosphérique ($p = 1,013.105$ Pa), les courbes (figure 2.3) passent par un minimum pour une distance interélectrode d'environ $10 \mu\text{m}$ ($p.d = 0,5$ à 5 environ). Cela signifie que, pour la production de décharges à pression atmosphérique dans des conditions d'utilisation usuelles ($d \geq 1$ mm), la tension de claquage est une fonction croissante de la distance interélectrode.

Les courbes de Paschen pour le SF_6 et l'azote sont représentées sur les figures 2.4, et 2.5. Ces courbes, bien que légèrement différent d'un gaz à l'autre, présentent toutes la même allure générale, à savoir : une courte décroissance ; un minimum de potentiel qui, pour la plupart des

gaz purs compris entre 100 et 500 Volts, ensuite une décroissance prolongée qui montre qu'au fur et à mesure que la densité augmente, il devient plus en plus difficile de rendre le gaz conducteur.

Aux faibles pressions, libre parcours moyen devenant du même ordre de grandeur que la distance interélectrode, l'influence de la cathode, et de façon générale des électrodes, devient prédominante au fur et à mesure que la densité (ou la pression à température constante) diminue, de ce fait, le nombre des collisions électroniques seront de moins en moins nombreuses. Alors si l'on veut $\bar{\alpha}$ et $\frac{\gamma}{\alpha}$ le plus élevés possibles, de manière à préserver la condition de criticité, il faut augmenter le nombre de collisions ionisantes. Dans ces conditions le champ E et la tension disruptive doivent être augmentés.

Dans le sens contraire, si la pression croît le nombre de collisions par unité de longueur augmente et le libre parcours moyen diminue ainsi que l'énergie moyenne que les électrons peuvent gagner par suite de l'action de champ électrique le long d'un libre parcours moyen. Ainsi, pour compenser ce comportement, la tension disruptive doit croître. Il s'ensuit que la courbe de Paschen doit avoir un minimum

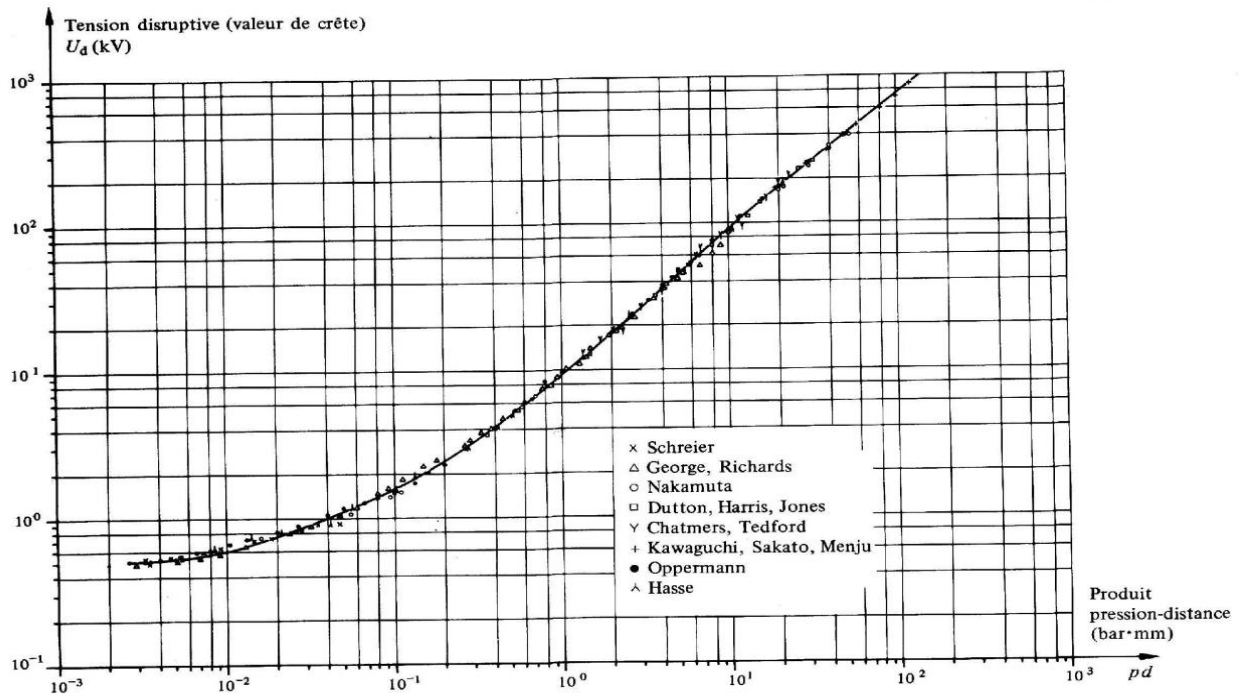


Figure 2.4: Courbe de Paschen pour SF₆ en échelles logarithmiques Températures 25°C [II.7]

Tension Disruptive (Valeur de crête)
 U_d (Volts)

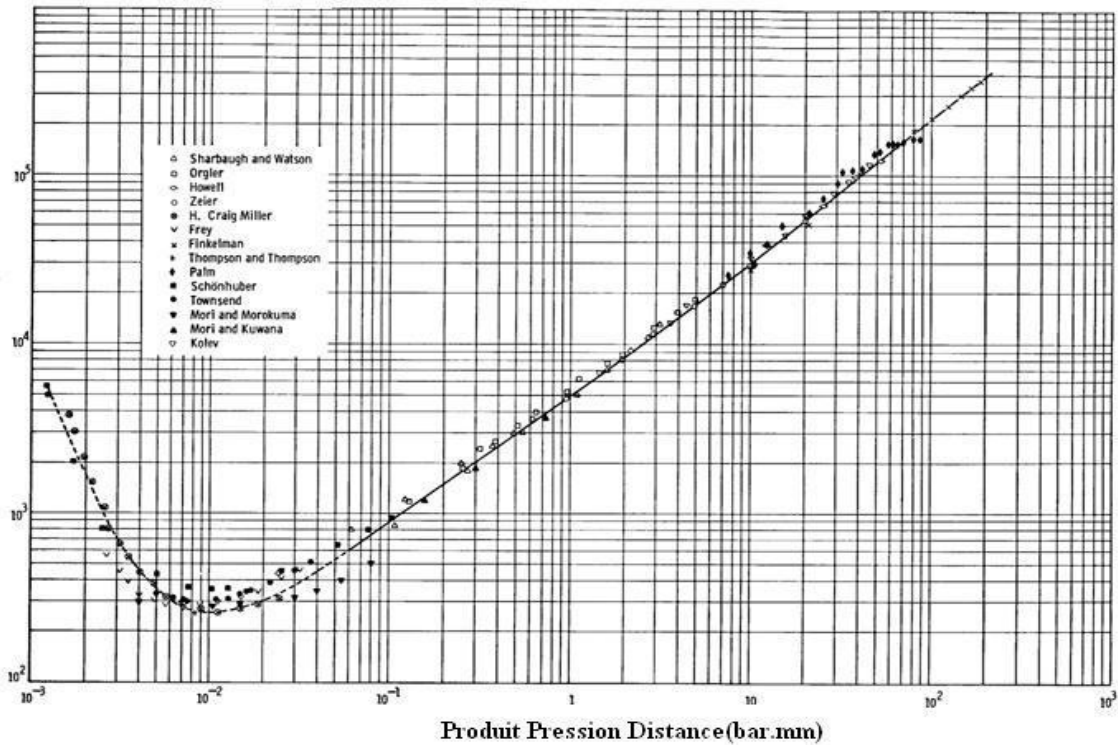


Figure 2.5: Courbe de Paschen pour l'azote en échelles logarithmiques Températures 25°C [II.7]

2.3.4 Déviations à la loi de Paschen

La théorie de Townsend fait intervenir tout l'intervalle interelectrode. Elle prévoit, à des pressions élevées, des temps de transition des charges entre les électrodes beaucoup plus grand que le temps de la formation de la décharge. Sur la figure 2.6, sont montrées les courbes déviées du SF_6 .

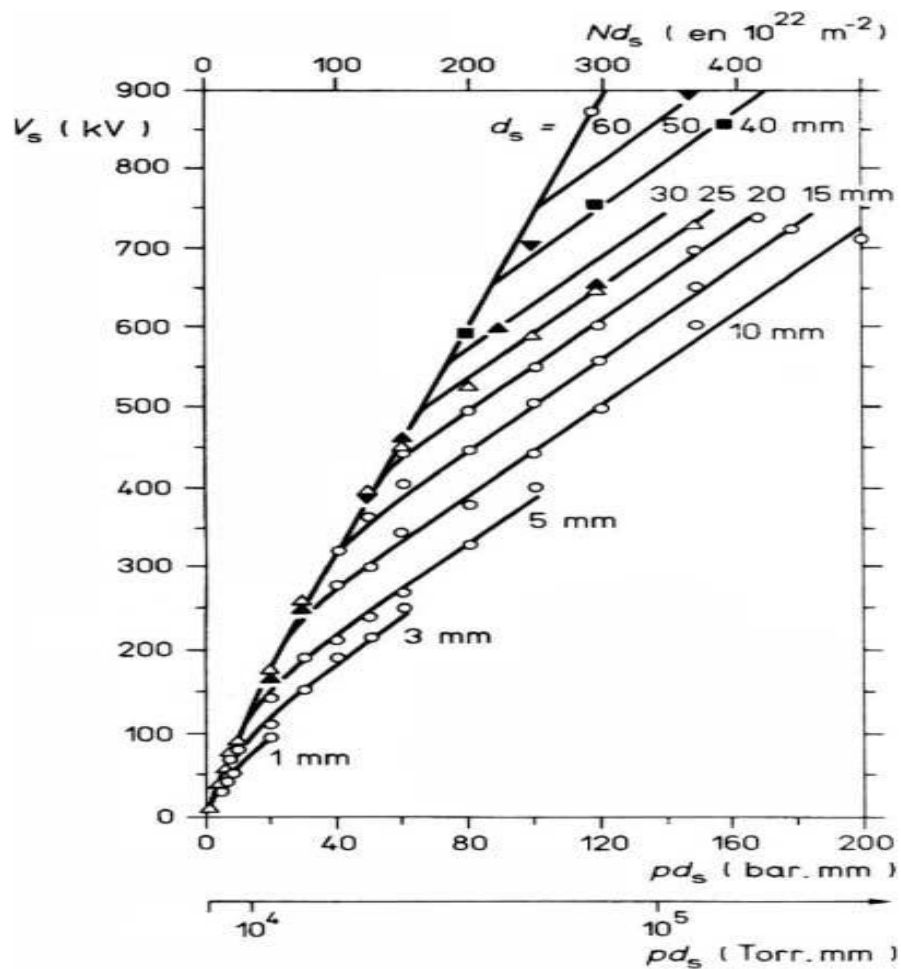


Figure 2.6 : Courbe de Paschen pour l'azote en échelles logarithmiques Températures 25°C [II.7]

Les déviations de paschen se produisent surtout aux pressions élevées ou très faible distance. Ces déviations sont souvent associées à la présence de champs électrique élevés (>100 KV/cm) qui sont comparables à ceux où se produit une émission d'électrons à la cathode, renforcée par les phénomènes de surface, sont vide poussé

2.3.5 Théorie du streamer

Townsend a proposé en 1915, un modèle de décharge qui s'applique aux systèmes pointe positive-plan. Dans ce modèle, un électron présent au voisinage de la pointe est accéléré vers celle-ci par le champ électrique. La présence de cet électron est expliquée par la photo ionisation d'une molécule du gaz par le rayonnement cosmique. Cet électron, accéléré vers la pointe, subit lors de son trajet des collisions ionisantes, produisant une avalanche d'électrons attirés eux aussi vers la pointe. Après leur collecte par l'anode, ils laissent dans leur sillage des cations qui forment

Une zone de charges d'espace positive (Figure II.3). Celle-ci modifie localement le champ électrique, prolongeant celui créé par l'anode dans l'espace interélectrode c'est la formation du streamer ou du dard (notation française) [II.8], [II.9], [II.10], [II.11]. Le mécanisme de propagation du streamer a été proposé par Loeb [II.5], Meek [II.13] et Raether [II.9]

Cette théorie du streamer introduit la notion de production d'électrons par photoionisation (les photons venant de la désexcitation, après 10^{-8} s environ, des molécules excitées par impact électronique [II.14], en amont de la charge d'espace, mais suffisamment proche de celle-ci pour le champ électrique et engendrent des avalanches secondaires provoquant la propagation du streamer (Figure 2.7). Après passage du streamer, il se forme entre celui-ci et l'anode un plasma neutre englobé par une zone de charge positive. Le champ de ce plasma est très faible (de l'ordre de 5 kV.cm^{-1}), favorisant alors les phénomènes d'attachement électronique par les espèces neutres. La chute du champ électrique dans le plasma créé par le passage du streamer, rend les vitesses de diffusion plus importantes que les vitesses de convection liées directement au champ électrique.

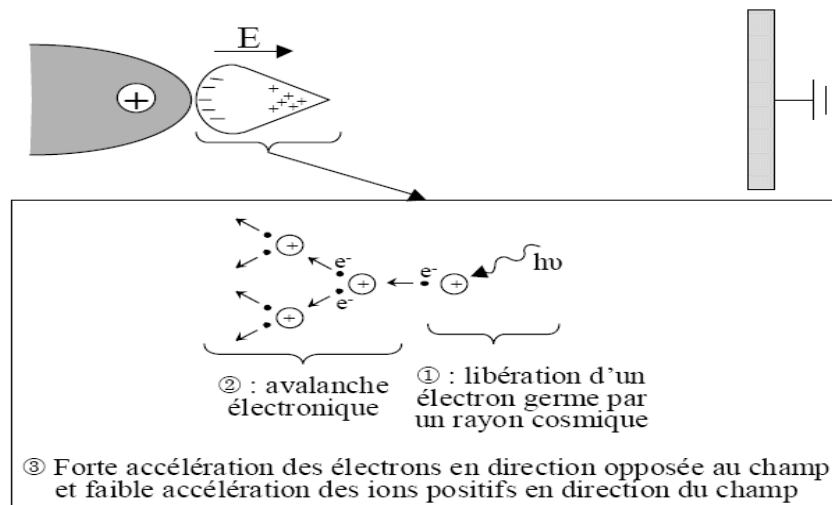
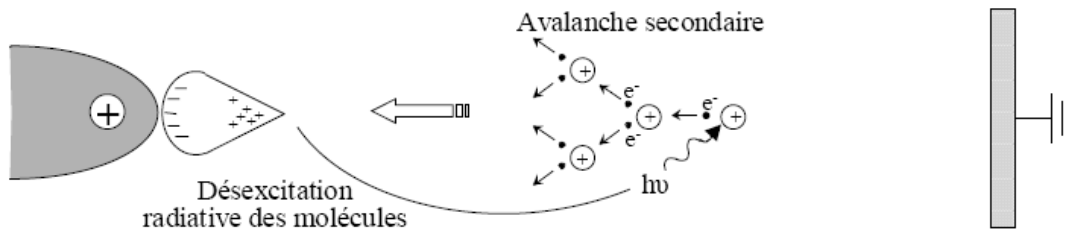


Figure 2.7: Mécanisme de développement du streamer proposé par Townsend



Propagation de proche en proche de la tête de streamer par création de nouvelles charges d'espace qui font progresser la zone à champ fort vers le plan

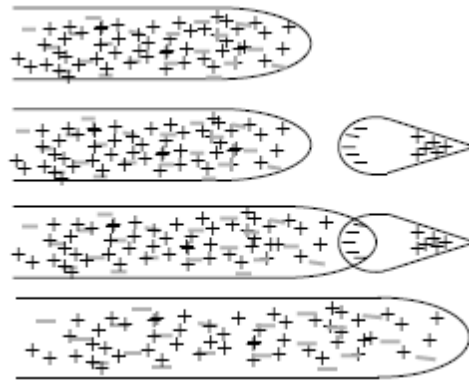


Figure 2.8: Mécanisme de propagation du streamer par photoionisation

Raether a établi un critère simple qui exprime que l'avalanche doit atteindre une taille critique L_c pour que le streamer puisse se développer. Dans le cas d'un champ non uniforme, si l'on désigne par E_s le champ à la surface de l'anode et E_c le champ critique du minimum d'ionisation correspondant à $(\alpha - \eta = 0)$, ce critère s'écrit :

$$\int_0^{L_c} (\alpha - \eta) dx = 1,5 \cdot 10^8 \text{ Électrons} \quad (2.4)$$

L'expression ci-dessus est obtenue du critère de claquage de Townsend et du critère de Meek. Ce dernier exprime que le champ de la charge d'espace E_p est du même ordre de grandeur que le champ E_s à la surface de l'anode.

2.4 Décharge couronne

Le terme décharge couronne désigne l'ensemble des phénomènes qui se produisent dans un milieu gazeux, situé entre deux électrodes dont l'une est active de faible rayon de courbure. Ce milieu est soumis à des champs électriques intenses mais non disruptifs et les électrodes peuvent former divers systèmes: pointe-plan, fil-plan, fil-cylindre, sphère-cylindre, etc.... Les phénomènes observés sont de nature optique, acoustique et électrique. Ils dépendent de trois classes de paramètres:

- a) Paramètres électriques (nature et polarité de la tension appliquée à l'électrode active).
- b) Paramètres géométriques (forme, dimension et distance interelectrode).
- c) Paramètres du gaz interelectrode (nature, pression, température et humidité).

La décharge couronne ne nécessite aucune source extérieure d'ionisation pour se développer, on dit que c'est une décharge auto-entretenu. Cependant, la présence d'un électron initial pour amorcer la décharge couronne est nécessaire. Cet électron germe doit se situer convenablement au voisinage de l'anode et son existence dans l'air est aléatoire. La décharge ne peut se former alors qu'après un certain temps correspondant, d'une part, à la création de l'électron germe, et d'autre part, à sa multiplication électronique ultérieure. Si une tension continue, suffisante pour créer une décharge couronne, est appliquée à l'électrode de faible rayon de courbure, l'espace interelectrode se répartie alors en deux zones:

- a) Une zone active, située près de l'électrode à forte courbure, dans laquelle le champ électrique est intense. Elle est caractérisée par l'ionisation poussée du gaz.
- b) Une zone de dérive, occupant l'espace interelectrode restant, dans laquelle le champ électrique est faible. Elle est caractérisée par une charge d'espace monopolaire dérivant vers la seconde électrode du système.

Les recherches antérieures ont conduit à distinguer deux processus de développement de la décharge. Le premier fait intervenir l'intervalle interelectrode tout entier, c'est le cas de la décharge de Townsend [II.15]. Le deuxième correspond à l'apparition de lueurs dans l'intervalle prenant l'aspect de filaments, c'est la formation du dard (streamer en notation anglo-saxon) [II.16].

2.4.1 Propriétés spécifiques aux décharges couronne

2.4.1.1 Caractéristique courant-tension et régimes de décharge

La caractérisation courant-tension de la décharge est obtenue en faisant croître la tension appliquée à l'intervalle gazeux. Différents régimes de décharge sont alors parcourus (Figure 2.9).

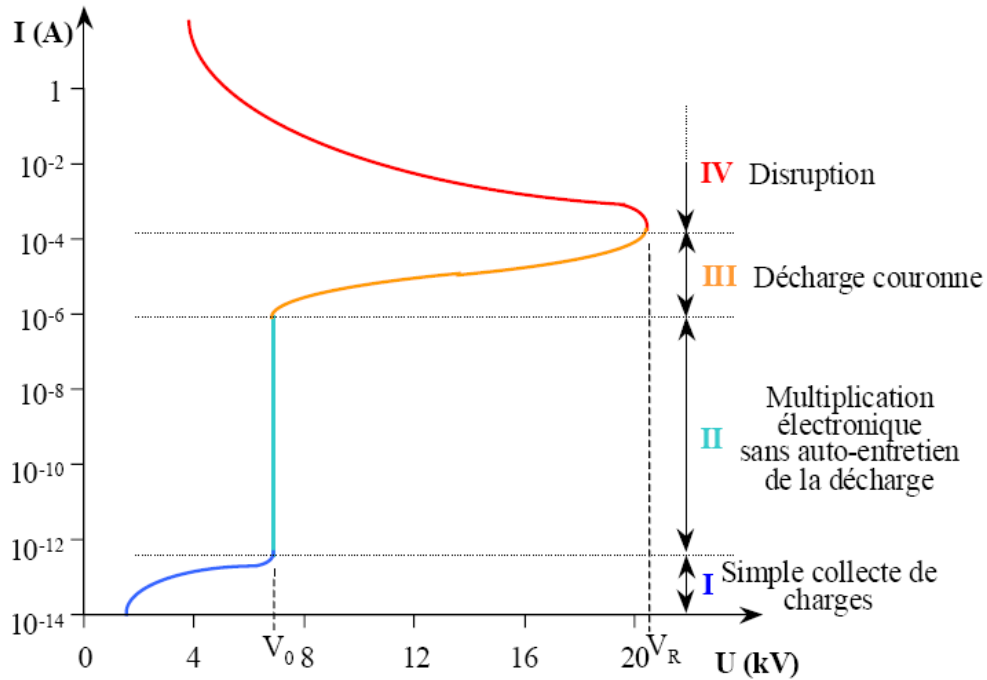


Figure 2.9: Caractéristique courant-tension d'une décharge couronne dans l'air à pression atmosphérique (configuration pointe-plan, distance interélectrode= 13 mm) [II.17]

Le premier régime (I) consiste en la collecte des espèces chargées présentes naturellement dans le milieu. Ces espèces sont produites par l'impact de rayons cosmiques ou de particules issues de la désintégration spontanée de certains atomes (radioactivité naturelle) sur le gaz présent dans l'espace interélectrode.

Le second régime (II), qui se produit à partir d'une tension seuil V_0 , traduit le déclenchement de l'avalanche électronique de Townsend mais ne répond pas encore au critère d'auto-entretien de la décharge. La décharge est qualifiée de non-autonome car elle dépend de processus d'ionisation externes pour donner naissance aux électrons germes. Pour ce régime, le courant augmente très fortement pour une très faible variation de la tension appliquée.

Pour **le troisième régime (III)**, l'ionisation du gaz par la décharge est suffisante pour ne plus nécessiter de phénomènes d'ionisation extérieurs : la décharge est auto-entretenu. C'est le régime de la décharge couronne, où apparaissent les dards (aussi appelés "streamers") responsables de la composante impulsionnelle du courant.

Au-delà d'une tension appliquée V_r , la température du canal augmente fortement, ce qui provoque la dilatation brutale du gaz. Le champ réduit E/N est alors suffisamment important pour que les phénomènes d'ionisation deviennent majoritaires par rapport à l'attachement électronique.

Le régime d'arc (4) est atteint, le milieu passe de l'état de plasma hors équilibre à l'état de plasma thermique.

2.4.1.2 Champ électrique et topographie de la décharge

La dissymétrie des électrodes entraîne la non-homogénéité du champ électrique. Parmi les configurations les plus étudiées, nous retrouvons les configurations pointe-plan. La distribution spatiale du champ électrique est donnée par la relation suivante: En géométrie pointe hyperbolique-plan :

$$E_s(x) = \frac{U_s}{\left(x + \frac{r}{2}\right) \text{Ln} \frac{2d+r}{r}} \quad (2.5)$$

Avec

$E(x)$: champ électrique ($V.m^{-1}$),

U : tension appliquée à la pointe

x : distance à la pointe sur l'axe pointe-plan (m)

r : courbure de la pointe à son extrémité (m)

d : distance interélectrode (m)

Des simulations ont été effectuées en configuration pointe-plan pour déterminer le champ électrique dans l'espace inter électrode avant ou après passage du streamer Figure 2.10 [II.18]

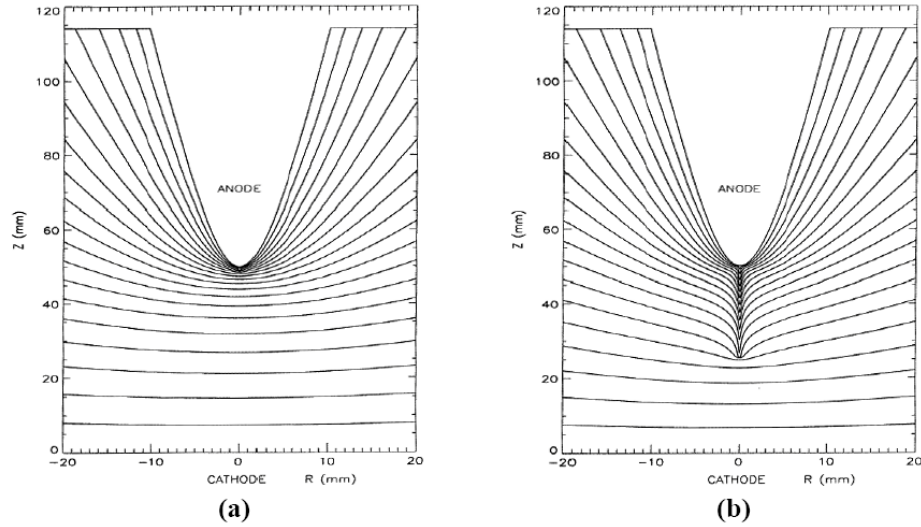


Figure 2.10: **(a)** Lignes isopotential avant le passage du streamer
(b) 10 μ s après passage du streamer [II.18]

Dans le cas d'une géométrie pointe-plan (figure 2.11), le champ intense qui règne à proximité de l'électrode active ($10^{-19} < E/N < 10^{-17} \text{ V.m}^{-2}$) permet l'ionisation du gaz. Les ions présents dans cette zone peuvent être positifs ou négatifs. Au-delà de cette région se trouve une zone de dérive ionique où le champ électrique est trop faible pour permettre les réactions d'ionisation ($10^{-21} < E/N < 10^{-19} \text{ V.m}^{-2}$). Des ions négatifs sont alors formés par attachement électronique. Ils sont ensuite entraînés le long des lignes de champ et réagissent avec les molécules rencontrées. La zone de dérive ionique, électriquement passive, est le siège d'une intense activité physico-chimique (dispersion des espèces par le vent électrique, réactions radicalaires...). A l'interface entre la zone d'ionisation et la zone de dérive ionique, la consommation d'électrons par attachement électronique égale leur libération par impact électronique : les coefficients α et η sont alors égaux.

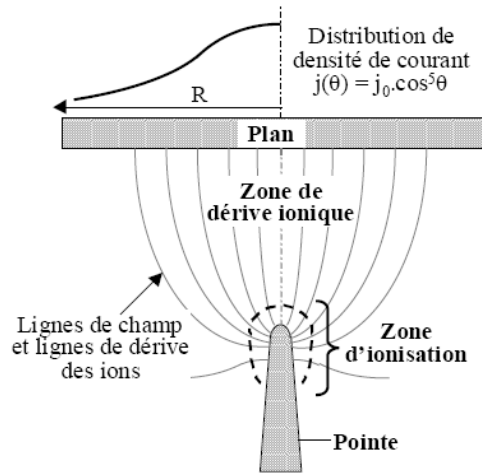


Figure 2.11: Topographie de la décharge couronne en configuration pointe plan [II.19]

2.5 Décharge électrique dans le mélange SF₆-N₂

L'association de N₂ et de SF₆ permet d'utiliser au maximum les qualités de chacun des gaz. En effet, le SF₆ est un gaz électronégatif dont les capacités d'attachement sont particulièrement efficaces. L'azote, quant à lui, a de bonnes qualités de ralentissement des électrons énergiques. Ainsi dans le mélange les électrons aptes à ioniser les atomes et à déclencher une avalanche électronique, sont ralentis par les molécules d'azote et arrivent dans des domaines d'énergie où le SF₆ a les capacités d'attachement les plus grandes.

Le coefficient d'ionisation effectif pour le mélange SF₆-N₂ est donné par l'expression suivante

$$\left(\frac{\bar{\alpha}}{P}\right)_{\text{lim}} = z \cdot (\bar{\alpha}/P)_{\text{SF}_6} + (1-z) \cdot (\alpha/P)_{\text{N}_2} \quad (2.6)$$

Avec z = le taux du SF₆ dans le mélange : $z = P(\text{SF}_6)/P$

$$\frac{\bar{\alpha}}{P} = z\beta \left[(E/P) - (E/P)_{\text{lim}} \right] + (1-z)A \exp\left(\frac{-B}{(E/P)}\right) \quad (2.7)$$

$(E/P)_{\text{lim}}$ diminue avec la diminution de la quantité de SF₆ dans le mélange et l'expression précédente prendra la forme suivante :

$$\left(\frac{\bar{\alpha}}{P}\right)_{\text{mél}} = \beta_m \left[(E/P) - (E/P)_{\text{lim}} \right] \quad (2.8)$$

Malik et Qureshi ont calculé $(E/P)_{\text{lim}}$ pour le mélange $\text{SF}_6\text{-N}_2$ en supposant que l'azote et le SF_6 ne réagissent pas avec les électrons de même énergie.

Cependant cette supposition n'est pas rigoureusement correcte. Kline et al. ont utilisé l'expression empirique suivante pour déterminer $(E/N)_{\text{lim}}$ et qui semble donner des résultats plutôt satisfaisant.

$$\left(\frac{E}{N}\right)_{\text{lim}} = \left(\frac{E}{N}\right)_{\text{SF}_6} \cdot (\% \text{SF}_6)^{0.18} \quad (2.9)$$

Où N est la densité.

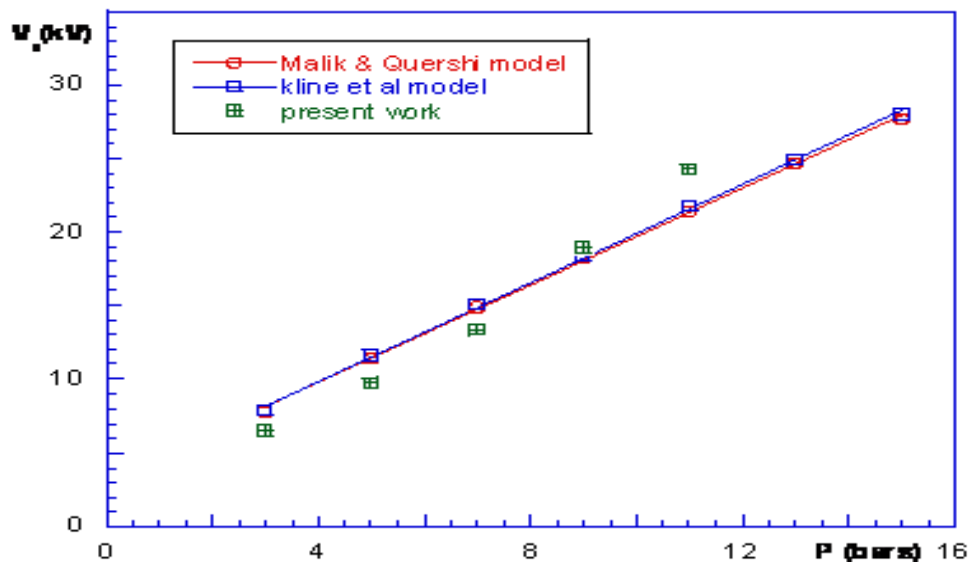


Figure 2.12: Comparaison entre les résultats expérimentaux de ce travail, les résultats trouvés par Malik et Qureshi et ceux déterminés en utilisant Kline et al modèle pour les pur SF_6 [II.17]

2.6 La conduction électrique dans les gaz isolants

Dans cette partie, nous présentons une Étude bibliographique des théories de transport des électrons et des ions et, des méthodes de détermination de leurs mobilités. Enfin, nous rappelons comment mesurer la vitesse des porteurs de charge soit par la méthode directe dite du "temps de vol" (utilisant des électrodes pointe-plan), soit par une méthode indirecte utilisant les caractéristiques courant-tension obtenues en géométrie pointe-plan

2.6.1 Théorie de la mobilité ionique de Langevin

Cette partie En 1905 Langevin [II.20] a publié sa première théorie sur la mobilité ionique basée sur la méthode des transferts de quantité de mouvement. Cette théorie s'applique à une région de (E/p) champ électrique faible (comme c'est le cas dans la zone de dérivé en configuration pointe-plan), elle tient compte l'effet de la force d'attraction entre une molécule et un ion de charge e , à une distance r , et également de la force de répulsion entre sphères rigides. Un ion attire une molécule neutre par les forces de polarisation dans le cas d'un gaz non-polaire, et par attraction directe des dipôles permanents dans le cas d'un gaz polaire. Cette force d'attraction f est donnée par l'expression suivante:

$$f = \frac{(\epsilon - 1)e_i^2}{2\pi N r^5} \quad (2.10)$$

Où ϵ représente la permittivité; e_i est la charge ionique; N la densité moléculaire et r la distance entre l'ion et la molécule. Cette expression est valable lorsque r est plus grande que la séparation des charges du dipôle. Une décroissance du libre parcours moyen des ions résulte de l'augmentation du taux des collisions, et c'est ainsi que le transfert de la quantité de mouvement s'établit entre les ions et les molécules même s'ils n'entrent pas en collision.

L'énergie du champ est négligeable par rapport à l'énergie thermique. Langevin a obtenu l'équation suivante pour la mobilité (μ):

$$\mu = \frac{e\lambda}{m\bar{v}} \quad (2.11)$$

λ représente le libre parcours moyen; \bar{v} est la vitesse thermique moyenne; e est la charge électrique de l'ion.

En prenant en considération les forces d'attraction, Langevin a formulé l'expression suivante pour le calcul de la mobilité ionique d'un gaz:

$$\mu = \frac{A(\lambda)}{\sqrt{\rho(\epsilon - 1)}} \sqrt{\frac{M + m}{m}} \quad (2.12)$$

Où ρ représente la masse volumique du gaz. A est une grandeur dépendante de λ et dont la variation est donnée par [2.2], ϵ est la constante diélectrique du gaz, m est la masse ionique et M est la masse de la molécule. Pour des valeurs relativement large de A ($A=0.75$), l'expression (2.12) prend la forme suivante: (cas limite de la sphère élastique)

$$\mu_e = \frac{0.75e_i}{D_{12}\sqrt{8\pi\rho}} \sqrt{\frac{M+m}{m}} \quad (2.13)$$

P est la pression en atmosphère, e_i est la charge ionique, D_{12} est la somme des rayons des molécules et d'ions.

Pour des valeurs de λ approchant zéro les forces de polarisation prédominent et pour $A = 0.5105$, l'expression (2.12) devient: (cas limite de polarisation)

$$\mu_p = \frac{0.5105}{\sqrt{\rho(\epsilon-1)}} \sqrt{\frac{M+m}{m}} \quad (2.14)$$

Dans l'expression (2-10) l'absence de la charge ionique, malgré que la force exercée sur l'ion à cause du champ E soit directement proportionnelle à la charge (e) s'explique par le fait que la perte de la quantité de mouvement est aussi proportionnelle a (e) et par conséquent (e) est simplifiée. L'indépendance à la variation de la température s'explique de la même manière, c'est à dire avec l'augmentation de la température la vitesse thermique des ions augmente et la perte de la quantité de mouvement diminue suffisamment pour éliminer l'effet de la température. L'équation de Langevin dans la limite de polarisation apparaît comme l'équation la plus appropriée pour la détermination de la mobilité ionique.

En utilisant l'approche quantique la théorie de Langevin évoluée prend la forme suivante:

$$\mu = \frac{35.9}{p\sqrt{(\alpha_p / a_0^3)}.M} \sqrt{\frac{m+M}{m}} \quad (2.15)$$

Où α_p est la polarisabilité (en unité atomique) ; a_0 est le rayon de Bohr.

2.6.2 Détermination de la mobilité par la méthode directe dite de temps de vol

L'estimation directe de la mobilité ionique consiste à mesurer le temps de transit des ions.

$$k = \frac{d^2}{vt} \quad (2.16)$$

v : est la différence de potentiel entre deux électrodes en parallèle.

t : est le temps de transit des ions.

d : est la distance inter-électrodes.

La conductivité naturelle du diélectrique est très faible et irrégulière et par conséquent pour mesurer la mobilité la densité des charges doit être augmentée ceci peut se faire par une excitation externe transitoire.

Les principales méthodes d'excitation utilisées sont les suivantes:

- Faisceau des rayons X qui peut créer des charges positives et négatives [II.21]
- Flash ultraviolet (UV) pour injection des électrons sur une photocathode. Ce procédé est facile mais ne peut mesurer que les porteurs de charges négatifs [II.22]
- Faisceau d'électrons de quelques KeV d'énergie où la densité des électrons est réglable.

Pour avoir une comparaison juste entre ces travaux, on a introduit le terme de mobilités réduites (μ_0) qui sont des valeurs mesurées pour une pression standard et une température standard et qui peuvent être déterminées par l'expression suivante :

$$\mu_0 = \frac{p}{760} \frac{273.16}{T} \mu \quad (2.17)$$

Où p est la pression du gaz exprimée en Torr et T est la température du gaz exprimée en degré Kelvin.

a) Mobilités des ions négatifs

La seule analyse de masse des porteurs de charges dans le SF_6 a été réalisée par Patterson [II.23]. Les mobilités les plus rapides sont attribuées aux ions SF_5^- ceux-ci ont été observés par Fleming et Rees [II.24] et mesurés par L. Patterson [II.23]. Les deux autres porteurs de charges trouvés par [II.24] ainsi que par Naidu et Prasad [II.25] correspondent aux ions SF_6^- et le premier cluster $SF_6^-(SF_6)$. Mobilité déterminée par Urquimo-Carmona [II.26] est attribuée au cluster $SF_6^-(SF_6)$ pour des valeurs de E/N pouvant atteindre la valeur critique de $3.6 \cdot 10^{-19} \text{ Vm}^2$. Il faut savoir que la valeur de $(E/N)_{lim}$ correspond à la condition critique de claquage où le coefficient d'ionisation et le coefficient d'attachement sont égaux ($\alpha = \eta$). Les mobilités trouvées par Crichton et Lee [II.27] sont proches de celles du SF_6^- . Les travaux d'Aschwanden [II.28] ont montré que pour des valeurs de E/N proche de $(E/N)_{lim}$ les mobilités des ions SF_6^- sont relativement élevées.

b) Mobilités des ions positifs

Les mobilités des différentes espèces ont un maximum pour un champ égal au champ critique $(E/N)_{\text{lim}}$. Au-dessous de cette valeur seulement deux résultats sont publiés, ceux de Fleming et al [II.24] et ceux d'Aschwanden [II.28]. Les mobilités trouvées sont proches de celles de l'ion SF_5^- ce qui laisse penser qu'il s'agit probablement de l'ion SF_5^+ dont la structure est inconnue. Cette attribution est consolidée par le fait qu'initialement l'ion SF_6^+ est formé avant de se dissocier rapidement en SF_5^+ . Au-dessus de $(E/N)_{\text{lim}}$, les résultats d'Aschwanden sont en parfaite concordance avec ceux de Urquijo-Carmona et de Teich et Sangi [II.29]. Par contre pour des valeurs très élevées de E/N des écarts importants sont observés. On suppose que les fortes mobilités mesurées précédemment sont dues essentiellement à la dissociation de SF_6^+ et les faibles mobilités peuvent être le résultat d'une dissociation d'un cluster positif. Jungblut et al [II.30] ont réalisé une expérimentation sans système d'identification d'espèces ni de polarité et ils sont arrivés à une conclusion dans laquelle ils admettent que les mobilités des ions positives et des ions négatifs sont presque égales pour un certain intervalle de E/N et pour une pression inférieure à 1 bar (Pas de formation de clusters).

Sur le tableau II-1, K. P. Brand et H. Jungblut [II.31] ont présenté un résumé sur la majorité des travaux concernant les différentes techniques de mesure de la mobilité utilisant la méthode directe. Il faut noter que les mobilités mentionnées correspondent aux mobilités réduites pour un champ très faible ou nul μ_{00} .

Création des porteurs de charges	Agent ionisant ou particule injectée	Variable mesurée	Polarité des ions	Type des ions	$\mu_{00} \cdot 10^{-15}$ m ² /Vs	Intervalle de E/N utilisé (10 ⁻¹⁹ V.m ²)	Pression (torr)
PEME pulsé (~1 ns) lampe Hg [II.31]	Photoélectrons	Temps de transit	négative	SF ₆ ⁻ (SF ₆)	4.5	0.07-0.74	0.88
Source des rayons α [II.21]	Particules α	Temps de transit	négative	SF ₆ ⁻	5.7	0.08	35
PEME PULV [II.32]	Photoélectrons	Temps de transit	négative	SF ₆ ⁻	5.7		0.34-0.69
Emission thermoïonique [II.24]	électrons	Temps de transit	négative	SF ₅ ⁺ ;SF ₅ ⁺	5.9 5.6	0.1-2.4	
Source α [II.24]	Particules α	Temps de transit	négative positive	SF ₆ ⁻ SF ₆ ⁻ (SF ₆)	5.4 4.7	0.1-2.4	
PEME pulsation de lumière répétée basse pression H-lampe arc [II.25]	Photoélectrons	Temps de transit	négative	SF ₆ ⁻ SF ₆ ⁻ (SF ₆)	5.6 4.7	0.3-2.4	1-10
Décharge pulsative répétée (~100cps) [II.24]	Electrons UV	Temps de transit	négative	SF ₅ ⁺ ;SF ₆ ⁻ SF ₆ ⁻ (SF ₆) SF ₆ ⁻ (SF ₆) ₂	5.94 5.42 4.7 4.20	0.0361.4	0.464

Emission thermoïonique [II.33]	Photoélectrons	Ion cyclotron résonance	négative	SF ₆ ⁻	5.42	0.3-4.5	7.10 ⁻⁴ -0.5
PEME-PUVL (10.15ns) [II.28]	Photoélectrons	Temps de transit	Positive négative			3.6-36 3.8-8	0.2-400
Emission du champ de la pointe cathode [II.27]	Electrons	Temps de transit	négative			1.9-3.5	122-3000
Pulsation (≤ 2.5 ns) accélérateur [II.30]	Bremsstrahlung	Temps de transit	indéfinie			0.3-3.7	50-500
PEME pulsation (~ 17ns) Laser [II.26]	Photoélectrons	Temps de transit	négative positive			0.08-6 1.9-60	0.07-193
PEME-PUVL(~10 ns) Laser (~ 3 ns) [II.28]	Photoélectrons	Temps de transit	negative positive			3.7-6 3.7-5.5	1-50

PEME: photo-effect at metal electrode

PUVL: pulsed UV flash Lamp

Tableau II-1. Mobilités ioniques dans le SF₆ mesurées par les méthodes directes [II.31].

2.6.3 Détermination de la mobilité à partir des courbes I-V en configuration pointe-plan

Nous nous intéressons exclusivement au cas d'une géométrie pointe-plan. Les modèles les plus appropriés sont celui de Coelho et Debeau [II.34] et le modèle de Sigmond [II.35]. Ces modèles décrivent et analysent les caractéristiques courant-tension afin d'en déduire les mobilités des porteurs de charges.

2.6.3.1 Modèle de Coelho et Debeau

a) Géométrie pointe -plan

Dans cette géométrie (figure 2.13), la région de champ élevé où la génération de charge se produit, est confinée près de la pointe, ce qui évite le phénomène de claquage. La plus grande partie de la distance pointe-plan d est par contre soumise à un champ suffisamment faible (car $d \gg r_p$) pour qu'aucune génération de charge n'y soit possible, seul le transport de la charge créée près de la pointe pourra s'y produire. Il est ainsi possible d'avoir accès au mécanisme de génération de charge et à celui de son transport. La distance sur laquelle le champ électrique est supérieur à une certaine valeur critique est d'autant plus faible (et en conséquence la région de création de charge est d'autant plus localisée) que le rayon de courbure de la pointe est plus petit. Cette distance est de l'ordre du rayon de courbure de la pointe. L'électrode pointe peut être assimilée à un hyperboloïde -plan.

La géométrie pointe-plan peut donc être décrite analytiquement dans un système de coordonnées curvilignes orthogonales, La géométrie de cette configuration est présentée sur la figure 2.13)

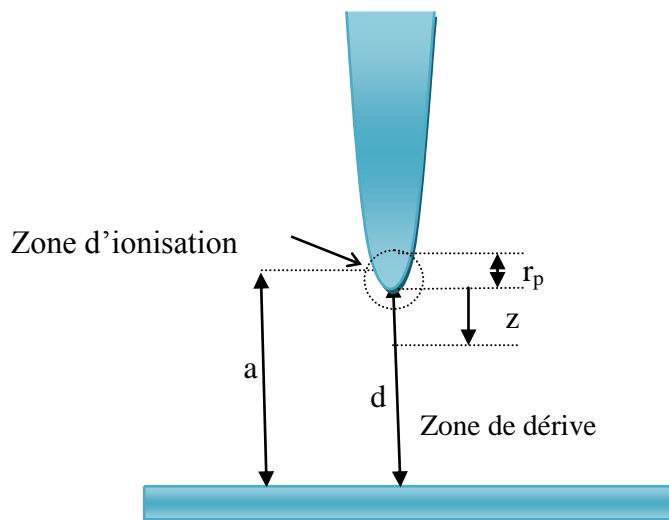


Figure 2.13: Géométrie d'une configuration hyperboloïde-plan

Une analyse mathématique de la configuration hyperboloïde-plan a permis à Coelho et Debeau d'établir une expression pour détermination de la mobilité en présence de charge d'espace à partir des courbes I-V. a cette effet ils ont résolu l'équation qui détermine la densité du courant j en présence de charge d'espace

En configuration pointe-plan la variation du courant moyen en fonction de la tension dépend de la mobilité des porteurs de charges. Le courant est limité par la charge d'espace.

Le champ en tout point d'espace est la somme du champ harmonique E_a et le champ du à la charge d'espace E_c . $E = E_a + E_c$

La distribution du champ électrique qui est gouverné par la charge d'espace est donnée par l'expression suivante:

$$E(z) = \left(1 - \left(\frac{z}{a}\right)^2\right)^{-1} \sqrt{\left(\frac{E_p r_p}{a}\right)^2 + \frac{4I}{a\mu} \left(2 - \frac{3z}{a} + \frac{z^3}{a^3}\right)} \quad (2.18)$$

r_p est le rayon de la pointe. Introduisant dans la formule (2.18) l'expression $(V-V_0)$:

$$E(z) = \frac{2V_0}{a \ln\left(\frac{4a}{r_p}\right)} \cdot \frac{1}{\left[1 - \left(\frac{z}{a}\right)^2\right]} \sqrt{1 + \left(\frac{V - V_0}{2V_0}\right)^2 \cdot \frac{\ln\frac{4a}{r_p}}{1.17} \left(2 - \frac{3z}{a} + \frac{z^3}{a^3}\right)} \quad (2.19)$$

Le premier terme de l'expression (2.19) représente la valeur du champ électrique sans charge d'espace (champ Laplacian):

$$E(z) = \frac{2V_0}{a \ln\left(\frac{4a}{r_p}\right)} \cdot \frac{1}{\left[1 - \left(\frac{z}{a}\right)^2\right]} \quad (2.20)$$

La valeur du champ électrique E_p proche de la pointe est:

$$E_p = \frac{2V}{r_p \ln \frac{4a}{r_p}} \quad (2.21)$$

Sur la figure 2.14, sont représentées les distributions du champ électrique avec et sans charges d'espace pour une configuration hyperboloïde-plan.

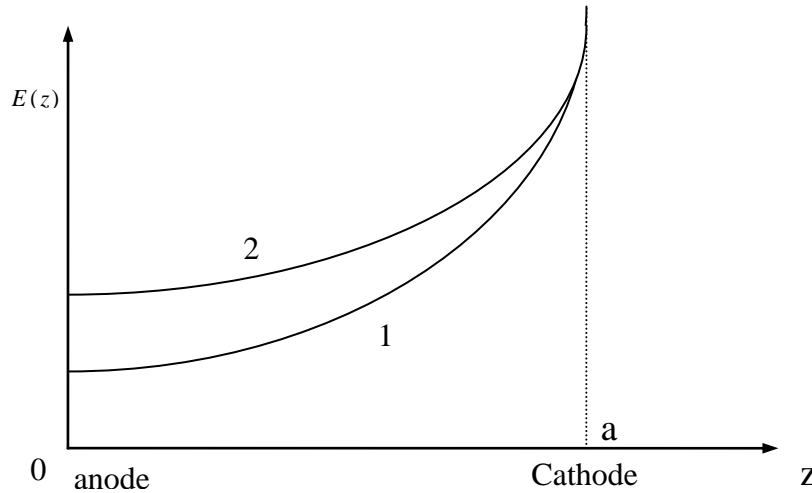


Figure 2.14: 1-Distribution du champ électrique pour une conduction sans charge d'espace
2- Distribution du champ électrique pour une conduction avec charge d'espace.

A partir de cette analyse ils obtiennent une relation donnant la variation du courant limité par la charge d'espace en fonction de la tension. Si le champ sur la pointe est nul l'équation du courant prend la forme suivante:

$$I_s = 4.02 \epsilon_0 \mu \frac{V^2}{a} \quad (2.22)$$

Par contre si le champ à la pointe est différent de zéro l'expression devient :

$$\sqrt{I_s} = 2.01 \sqrt{\frac{\epsilon_0 \mu}{a}} (V - V_0) \quad (2.23)$$

Il en résulte qu'à partir de la pente des courbes $\sqrt{I} = f(V)$ on peut déterminer la mobilité des porteurs de charges. Il faut noter que dans nos conditions $r_p \ll d$ et donc $d = a$.

Pour simplifier les calculs Coelho et Debeau ont considéré une distribution rectangulaire de la densité du courant sur l'électrode plane (anode), en estimant que l'erreur avec la distribution de Warburg [II.36] est négligeable.

b) Distribution de Warburg [II.36]

Comme on peut le voir sur la figure 2.15, Warburg a présenté la distribution de la densité du courant $j(\Theta)$ sur un plan dans une configuration pointe-plan.

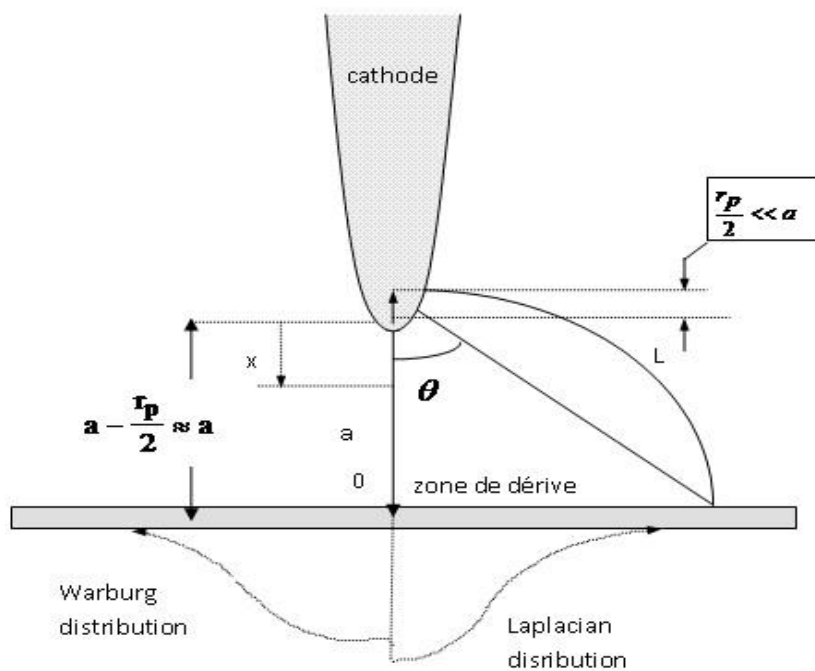


Figure 2.15: *Distribution de la densité de courant J proposé par Warburg [II-16].*

Il a formulé la distribution de la densité du courant $j(\Theta)$ sous forme de l'équation suivante:

$$j(\theta) = j(0) \cos^m \theta \quad (2.24)$$

$m = 4.82$ pour le cas d'une décharge couronne positive, $m = 4.65$ pour le cas d'une décharge couronne négative et $\theta \leq 60^\circ$

Cette estimation de la distribution de la densité de courant, a donné des résultats satisfaisants pour des domaines étendus des courants de décharges couronne.

2.6.3.2 Modèle de Sigmond

L'approximation de Sigmond [II.33], pour une décharge couronne unipolaire (une décharge unipolaire définie comme étant celle qui est conformée par des porteurs de la même polarité) saturée par la charge d'espace, est fondée sur la formule de la densité de charge unipolaire de saturée par la charge d'espace, est fondée sur la formule de la densité de charge unipolaire de dérive [II.35].

$$\frac{1}{\rho(\mathbf{t})} - \frac{1}{\rho(\mathbf{t}_0)} = \frac{\mu}{\epsilon_0}(\mathbf{t}) \quad (2.25)$$

La supposition que le flux ionique unipolaire divergent tend à égaliser les variations du champ électrique tout le long de ces mêmes lignes du champ. Pour une densité ionique initiale ρ_0 suffisamment grande l'équation précédente devient:

$$\rho_0(\mathbf{t}) \approx \frac{\epsilon_0}{\mu \mathbf{t}} \quad (2.26)$$

Sur l'électrode passive (plane), la densité des charges est:

$$\rho_p \approx \frac{\epsilon_0}{\mu T} \quad (2.27)$$

T est le temps de transit de la charge d'espace sur une ligne de champ L reliant la pointe au plan. $L = d\sqrt{1 + \tan^2 \theta}$. Ou θ est l'angle que fait L par rapport à l'axe Pointe plan. En prenant l'expression du champ électrique $E = \frac{U}{L}$, le temps de transit $T \approx \frac{L}{\mu E} \approx \frac{L^2}{\mu U}$

Ainsi, $\rho_{\text{plan}} = \rho(T) = \epsilon_0 \frac{E}{L}$; $J_{\text{plan}} = \mu E \rho_{\text{plan}} = \mu \epsilon \frac{E^2}{L}$ On a donc:

$$J_{\text{plan}} = \mu \epsilon \frac{U^2}{L^3} \quad (2.28)$$

Sigmond obtint l'équation exprimant la densité du courant limité par la charge d'espace:

$$J(\theta) = \mu E r_p = \mu \epsilon_0 \frac{U^2}{L^3} \approx \frac{\mu \epsilon_0 U^2}{d^3} (1 + \tan^2 \theta)^{3/2} \quad (2.29)$$

La valeur du courant total est déduit en faisant l'intégration de la densité du courant sur tout le plan $J(\theta)$. Pour cela il est indispensable de connaître la répartition du courant sur ce plan que Sigmond admet être identique à celle décrite empiriquement par Warburg (figure 2.15) pour des

valeurs de $\theta = 0^0$ à $\theta = 60^0$, l'expression du courant limité par la charge d'espace et appelé courant de saturation et il est donné par la relation suivante:

$$I_s = 2\mu \varepsilon_0 \frac{V^2}{d} \quad (2.30)$$

Cette formule présente un intérêt particulier parce qu'elle nous donne le limite supérieure absolue de courant couronne unipolaire qui peut être obtenu pour une tension U déterminée. En autres mots, si on mesure, pour les ions, on un rapport **courant/ (tension)²** pour une décharge couronne au delà de cette limite de courant, on pourrait être absolument certain : soit, qu'une partie des ces ions ne sont que des électrons plus rapides, et/ou soit qu'une partie de notre courant est bipolaire

2.7 . Mobilité ionique dans les mélanges gazeux

La loi de Blanc [II.36], pour la détermination de la mobilité des mélanges est donnée par la relation suivante:

$$\frac{1}{\mu_{\text{mixture}}} = \frac{x_1}{\mu_1} + \frac{x_2}{\mu_2} \quad (2.31)$$

Ou μ_1 et μ_2 sont les mobilités des ions dans les gaz purs et x_1 et x_2 sont respectivement les fractions molaires du gaz 1 et 2 dans le mélange.

$$x_1 = \frac{N_1}{N_1 + N_2} \quad x_2 = \frac{N_2}{N_1 + N_2} = 1 - x_1$$

L'équation suivante est par conséquent obtenue pour le mélange gazeux SF₆-N₂:

$$\mu_{(\text{mélange})} = \frac{1}{\frac{x_{\text{SF}_6}}{\mu_{\text{SF}_6}} + \frac{x_{\text{N}_2}}{\mu_{\text{N}_2}}} = \frac{\mu_{\text{SF}_6}}{\left[1 - \% \text{N}_2 \left(1 - \sqrt{\frac{\alpha_{\text{N}_2} \tilde{M}_{\text{N}_2}}{\alpha_{\text{SF}_6} \tilde{M}_{\text{SF}_6}}} \right) \right]} \quad (2.32)$$

$\alpha_{\text{SF}_6} = 43,5$; $\alpha_{\text{N}_2} = 11,8$ Sont, respectivement les polarisabilités du SF₆ [II.23] et N₂ [II.25]

$$\tilde{M}_{\text{SF}_6} = \frac{M_{\text{SF}_6}^- * M_{\text{SF}_6}}{M_{\text{SF}_6}^- + M_{\text{SF}_6}} \quad \text{et} \quad \tilde{M}_{\text{N}_2} = \frac{M_{\text{SF}_6}^- * M_{\text{N}_2}}{M_{\text{SF}_6}^- + M_{\text{N}_2}}$$

% N₂ est la fraction molaire de N₂ dans le mélange.

2.8 Conclusion

La détermination de la mobilité des porteurs de charges peut être déterminée par deux méthodes. La première consiste à mesurer le temps de vols des ions du gaz et la deuxième dite méthode indirecte basée sur la mesure des caractéristiques courant-tension de la décharge électrique. Dans le présent travail, la méthode de Sigmond a été choisie pour déterminer les mobilités ioniques des ions de SF₆ et du mélange SF₆-N₂.

Bibliographie

CHAPITRE 2

- [II.1] M. Amokrane Handala ; »Etude de la décharge sous tension alternative 50HZ- effet sur un interface isolante air/solide » ; thèse de doctorat d'état en électrotechnique.(U.M.M.T.O),Décembre 2007 .
- [II.2] Ibrahima Ndiaye ;Initiation et développement des décharges couronnes sur une surface de glace ; mémoire présenté à université du Québec à Chicoutimi comme exigence partielle de la maîtrise en ingénierie.mai 2003.
- [II.3] Mounir Hairour ; Etude diélectrique d'une isolation hybride gaz-solide pour appareillage haute tension ; thèse de doctorat (université de Montpellier II ; octobre 2007.
- [II.4] Fletcher R.C., 1949, Phys. Rev., Vol. 76, p. 1501.
- [II.5] Gary C., 1984, Les propriétés diélectriques de l'air et les très hautes tensions, Editions Eyrolles.
- [II.6] Gas discharge physics, Y.P. Raizer, ed. Springer-Verlag, **1991**.
- [II.7] T. W. Dakin, G. Luxa, G. Oppermann, J. Virgreux, G. Wind, and H. Winkelkemper, "Breakdown of gases in uniform fields. Paschen curves for nitrogen, air and sulfur hexafluoride," *Electra*, vol. 32, p. 61–82, 1974.
- [II.8] Meek J.M., 1940, Phys. Rev., Vol. 57, p. 722 - 728.
- [II.9] Raether H., 1941, Zeit. Phys., Vol. 117, p. 394-524.
- [II.10] Loeb L.B., Kip A.F., 1939, J. Appl. Phys., Vol.10, p.142.
- [II.11] Loeb L.B., 1948, Rev. Mod. Phys., Vol. 20, p. 151.
- [II.12] L.B. Loeb & A.F. Kip Electrical discharge in air at atmospheric pressure, *J. Appl. Phys.*, **1939**, vol.10, p. 142.
- [II.13] J.M. Meek, A theory of spark discharge, *Phys. Rev.*, **1940**, vol.57, p.722.
- [II.14] Les décharges électriques continues et transitoires induites par streamer sous pression atmosphérique pour la destruction des composés organiques volatils, Thèse de Z. Machala, Université Paris XI, **2000**.

- [II.15] Townsend J.M., 1914, Electricity in gases, Oxford University press.
- [II.16] Meek J.M., 1940, Phys. Rev., Vol. 57, p. 722 - 728.
- [II.17] Corona discharge physics and applications, R.S. Sigmond & M. Goldman, Electrical.
- [II.18] R Morrow and J.J. Lowke, Streamer propagation in air, , *J. Phys. D: Appl. Phys.*, **1997**, vol.30, pp. 614-627.
- [II.19] R.S. Sigmond & M. Goldman, Electrical breakdown and discharges in gases - Part B, NATO ASI series B89-B, **1989**.
- [II.20] Langevin P., " une formule fondamentale de la théorie cinétique ". *Ann. Chim. Phys., Série 8*, Vol. 5, pp. 719-720, 1969.
- [II.21] H. R. Hassé "Langevin's theory of ionic mobility" *Phil. Mag.* Vol 1, pp. 139-160, 1926.
- [II.22] E.W. MacDaniel and M. R. C. Dowel, *Phys. Rev.* 114, pp. 1028, 1959.
- [II.23] P. L. Patterson, "Mobilities of negative ions in SF₆", *J. Chem. Phys.* 53, pp. 696, 1970.
- [II.24] I. A. Fleming and Rees, *J. Phys. B* 2, 777, 1969.
- [II.25] M. S. Naidu and A. N. Prasad, *J. Phys. D* 3, 951, 1970.
- [II.26] J. de Urquijo-Carmona, Ph.D. Thesis, Victoria University of Manchester, UMIST. 1980.
- [II.27] B. H. Crichton and Dong-in Lee, *Proceeding of the Fifth International Conference on Gas Discharges*, Liverpool, (IEE, London), pp. 254-255. 1978.
- [II.28] T. Aschwanden, Faculty of Electrical Engineering, ETH. Zurich (Private communication, 1982).
- [II.29] T. H. Teich and B. Sangi, *Proc of the First Intern. Symp. on High Voltage Engineering.* pp. 391-394, (Munich, West Germany, 1972).
- [II.30] H. Jungblut, Ph.D. Thesis Freie Universitat, Berlin, 1981.
- [II.31] K. P. Brand and H. Jungblut "The interaction potentials of SF₆ ions in SF₆ parent gas determined from mobility data", *J. Chem. Phys.* 78(4), pp 1999-2007, 1983.
- [II.32] R. Coelho & J. Debeau "Properties of tip-plane configuration", *J. Phys. D: Appl. Phys.* Vol. 4, 1971.
- [II.33] R. S. Sigmond "Simple approximate treatment of unipolar space-charge. dominated Coronas. The Warburg law and saturation current", *J. Appl. Phys.* 53(2), 1982.
- [II.34] Waters R.T., Richard T. E. S. and Stark W.B., "The structure of impulse corona in a rod-plane gap I. The positive corona" , *Proc. Roy. Soc. London*, A315, pp. 1-25, pp. 1-25, 1970.
- [II.35] E. Warburg *Handbook, Derphysik Springer Berlin*, Vol. 14, pp. 154-155, 1927.

Chapitre 3

Mobilités des Porteurs de charges

3.1 Introduction

Les données expérimentales utilisées ont été obtenues précédemment au laboratoire d'électrostatique et des matériaux diélectriques LEMD du centre national de la recherche scientifique CNRS de Grenoble en France. La base de données consiste des caractéristiques courant-tension de la décharge couronne en configuration pointe-plan pour le SF₆ et le mélange SF₆-N₂. Les pressions utilisées varient entre 2 et 15 bars, les mesures sont effectuées pour la polarité négative et positive.

3.2 Technique expérimentale

Le schéma du dispositif a été présenté dans la figure 3.1 [III.1]. La cellule d'étude en acier inoxydable, contient deux électrodes, l'une plane, en acier inoxydable ayant un rayon de 12 mm connectée à un électromètre afin de mesurer le courant collecté, et l'autre dite de pointe fabriquée en tungstène par la technique d'électrolyse et possédant un rayon de courbure de l'ordre de quelques micromètres ; la pointe est connectée à la source de haute tension à courant continu (dont la tension peut atteindre jusqu'à 20 KV). La cellule est nettoyée et placée sous vide à l'aide d'une pompe primaire à une pression de l'ordre de $5 \cdot 10^{-2}$ Pa, avant l'introduction du gaz. Le SF₆ délivré est d'une pureté de 99.97% et l'azote utilisé est N60. Le mélange est réalisé par la méthode partielle avec l'introduction du SF₆ avant l'N₂. Les mesures sont réalisées pour des pressions élevées variant de 2 à 15 bars et des distances inter-électrodes variant de 4 à 10 mm.

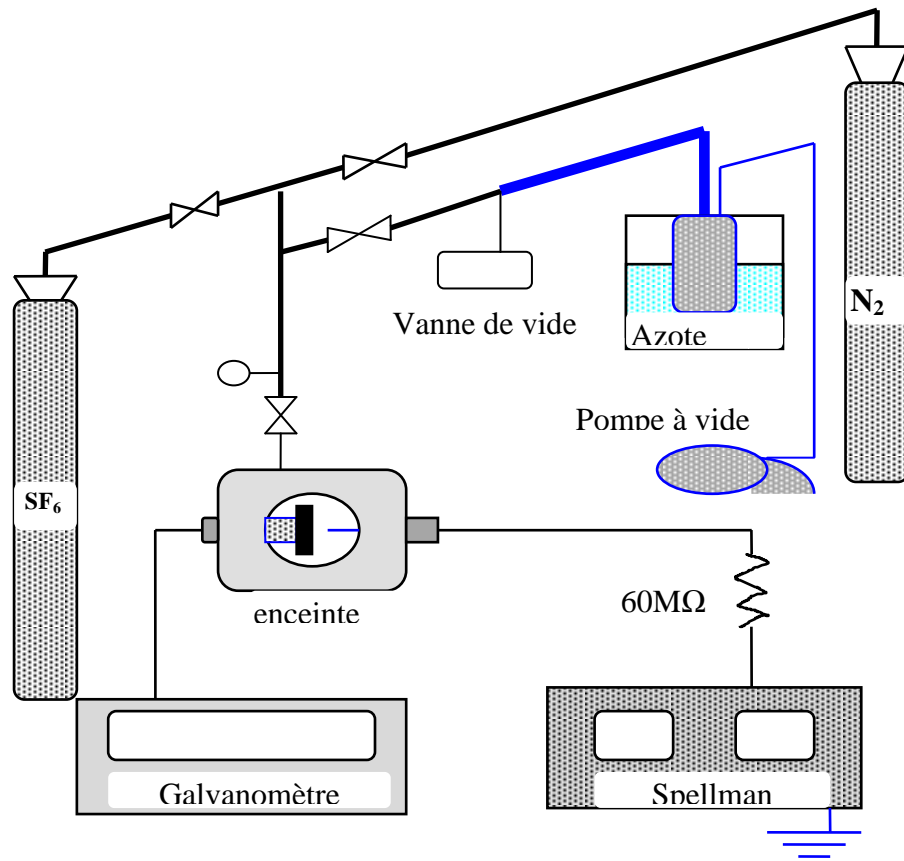


Figure 3.1: Dispositif de mesure des courant-tension ($I=f(U)$) pour la détermination des mobilités et des tensions seuils

3.2.1 Caractéristiques $I=f(U)$ et $\sqrt{I}=f(U)$ des décharges couronnes

Les caractéristiques $I=f(U)$ sont mesurées pour des pressions variant de 2 à 14 bars en polarité négative et en polarité positive pour le SF₆ pur (figure 3.1).

Sur la figure 3.2, sont représentées les caractéristiques $\sqrt{I}=f(U)$ pour le SF₆ à 100 % en polarité négative.

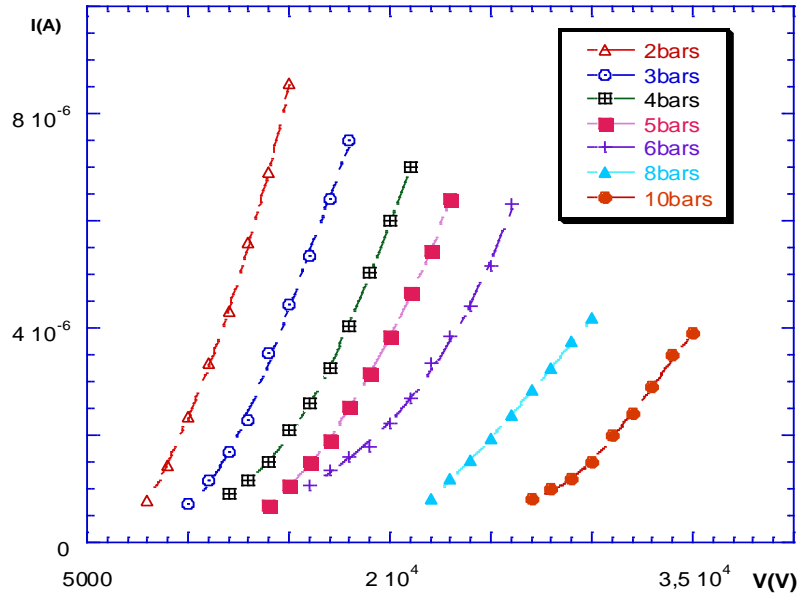


Figure 3.2:Caractéristiques $I(V)$ pour le SF_6 à 100 % à différentes pressions, en polarité négative

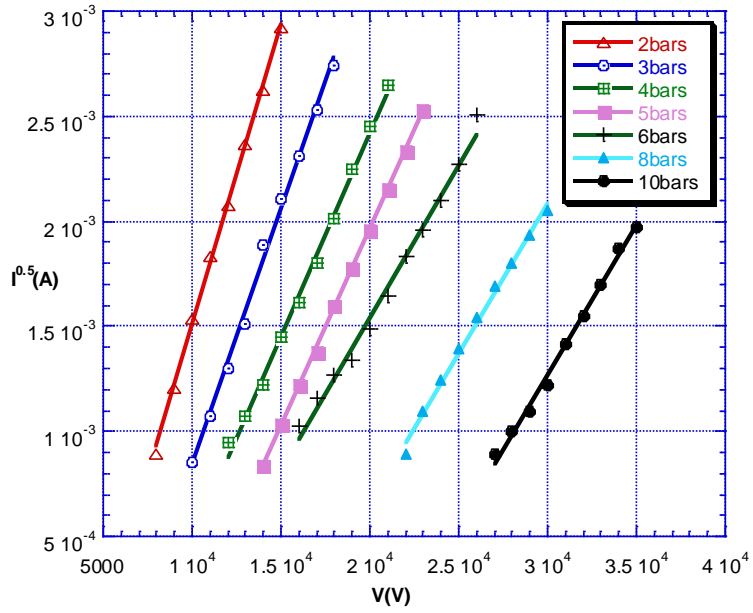


Figure 3.3:Caractéristiques $\sqrt{I}=f(U)$ pour le SF_6 à 100 % pour différentes pressions, en polarité négative.

Les caractéristiques $I=f(U)$ sont mesurées pour des pressions variant de 2 à 14 bars en polarité négative pour le mélange SF_6-N_2 à 10% de SF_6 (figure 3.3).

Sur la figure 3.4, sont représentées les caractéristiques $\sqrt{I}=f(U)$ pour mélange SF_6-N_2 à 10% de SF_6 en polarité négative.

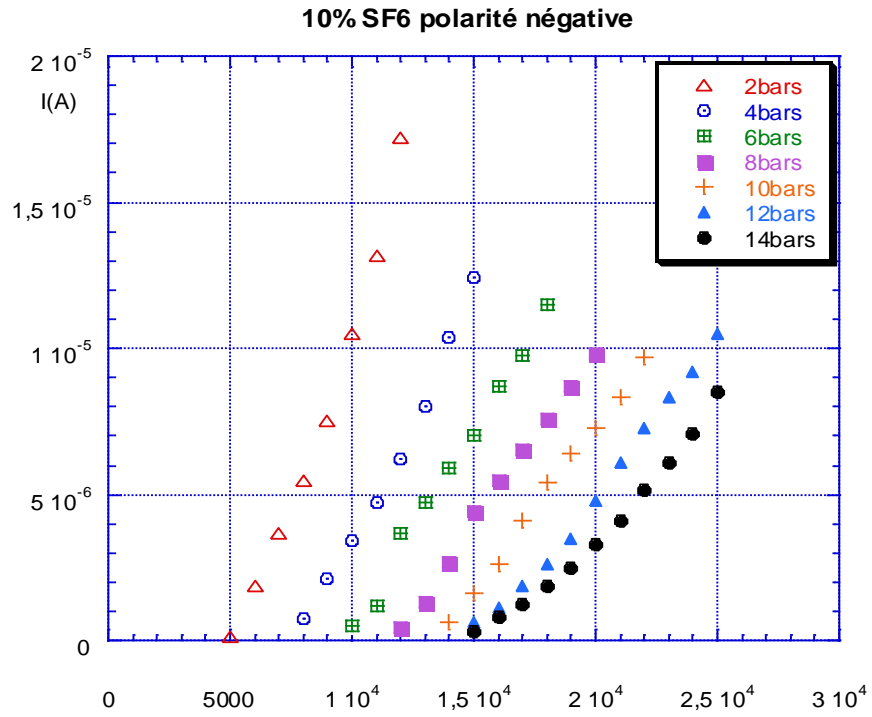


Figure 3.4: Caractéristiques $I(V)$ pour le 10% SF₆ / 90% N₂ à différentes pressions, en polarité négative.

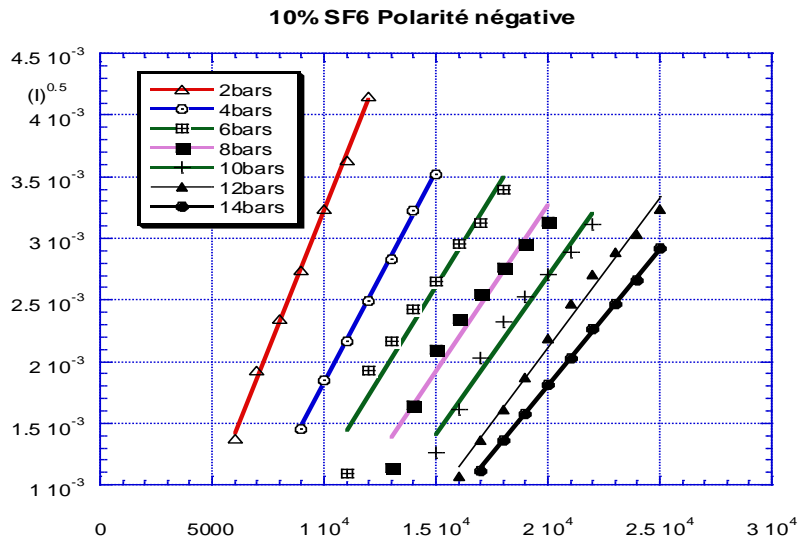


Figure 3.5: Caractéristiques $\sqrt{I} = f(U)$ pour le 10% SF₆ / 90% N₂ à différentes pressions, en polarité négative.

Dans Les figures 3.2, et figure 3. 4, présentent les courbes $I(V)$ pour différentes valeurs de pressions du gaz pour le SF_6 à 100% et pour le mélange 10% SF_6 / 90% N_2 respectivement. On peut constater qu'une certaine dépendance existe entre le courant et la densité du gaz, et que le courant mesuré reste toujours inférieur au courant de saturation de Sigmond on a donc un courant unipolaire.

Sur les figure 3.3et la figure 3.5, les courbes de $\sqrt{I}=f(U)$ sont tracées pour le SF_6 et SF_6-N_2 . A partir de ces courbes linéaires on peut déterminer les valeurs des pentes nécessaires pour le calcul de la mobilité.

3.2.2 Mobilité des porteurs de charges

Comme on peut le voir sur la figure 3.6, la mobilité μ dans cette gamme de pression est inversement proportionnelle à la densité du gaz (N). Pour le SF_6 pur le modèle de Sigmond donne des valeurs très proches à celles calculées en utilisant la théorie de Langevin.

Par conséquent, le modèle le plus approprié pour la prédiction de la mobilité ionique avec des mesures directes à SF_6 est le modèle R.S Sigmond. Les seuls valeurs trouvées dans la littérature sont celles de Schmidt [III.3] prélevées dans des conditions presque similaires aux nôtres et sont légèrement inférieures aux valeurs déterminées dans le présent travail.

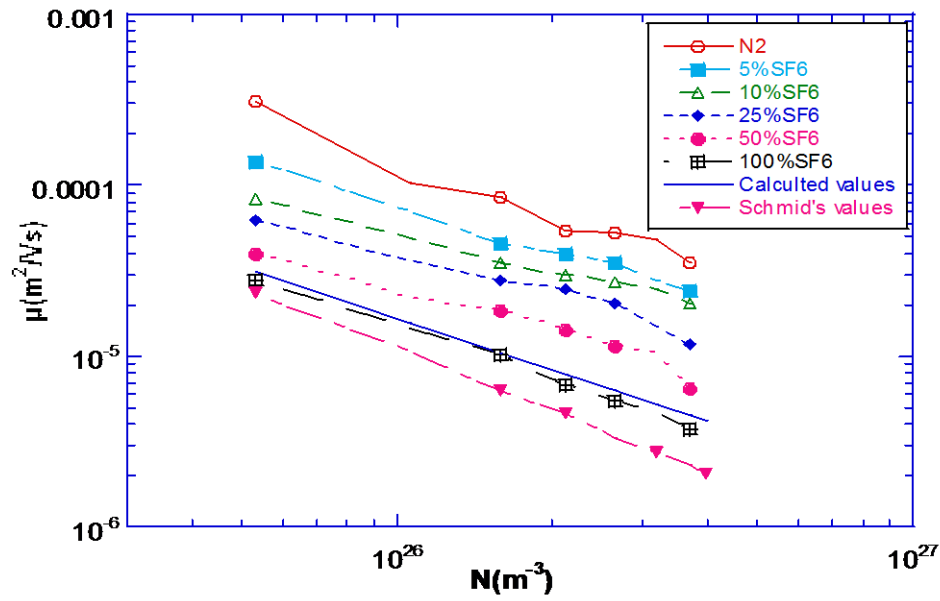


Figure 3.6: Variation de la mobilité en fonction de la densité du gaz (N).

Les courbes de la mobilité pour différentes densités de SF_6 dans les mélanges de gaz SF_6-N_2 en fonction de pourcentage de l'azote N_2 sont présentées sur la figure 3.7. Comme il peut être démontré, il y a une augmentation nette des mobilités avec la diminution du pourcentage du SF_6 dans le mélange gazeux.

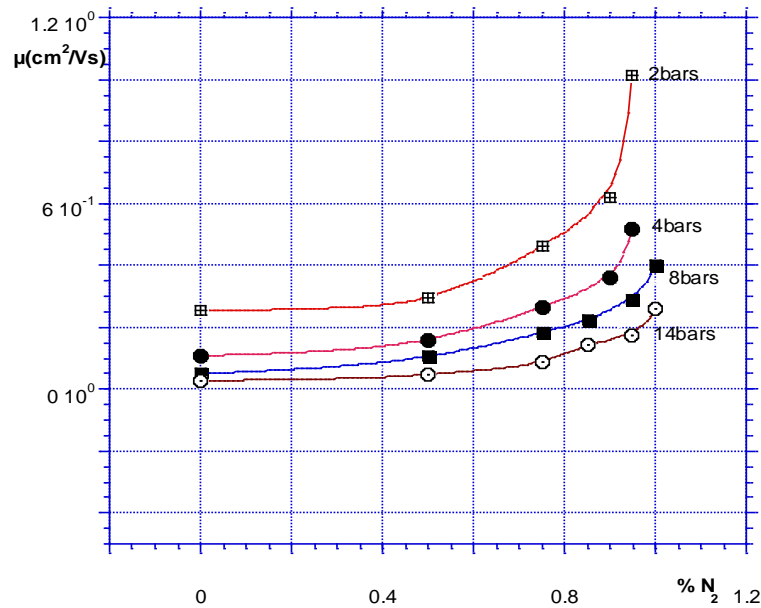


Figure 3.7: Mobilités en fonction du pourcentage de l'azote N_2 dans le mélange SF_6-N_2

Sur la figure 3.8, montre l'influence du rayon de courbure de la pointe sur les mobilités des porteurs de charges, pour le mélange SF₆-N₂ à 10% de SF₆.

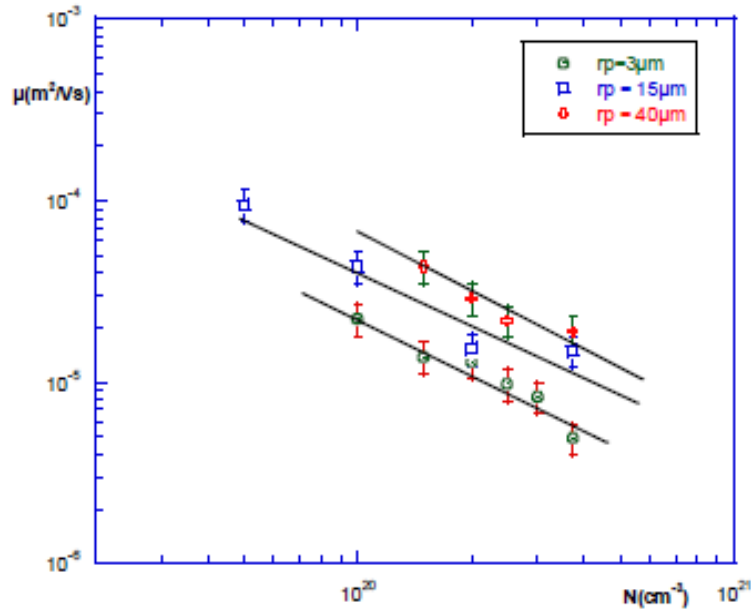


Figure 3.8: Influence du rayon de courbure de la pointe sur les mobilités des porteurs de charges, pour le mélange SF₆-N₂ à 10% de SF₆ en polarité négative

3.3 Conclusion

La variation des mobilités à pressions élevées pour le SF₆ pur est inversement proportionnelle à la densité du gaz (N^{-1}). Il est admis que l'ion SF₆⁻ est prédominant mais à pressions élevées la formation des ions plus lourds comme les clusters est très probable ce qui explique la différence observée entre la courbe théorique de Langevin et celle mesurée par Schmidt et le présent travail. Dans les mélanges gazeux SF₆-N₂, les mobilités augmentent avec modération en fonction du pourcentage du SF₆ dans le gaz jusqu'à une certaine limite du taux de SF₆ au delà duquel on assiste à une très forte augmentation des mobilités. On pense que ce phénomène est directement lié au changement de la nature des porteurs de charges, les ions O₂⁻ qui sont nettement plus légers, prennent le dessus sur les ions SF₆⁻.

Bibliographie

- [III.1] Lemzadmi A, Thèse Doctorat, Université de Annaba, 2006.
- [III.2] W. F. Schmidt and H. Jungbut “Ion mobility and recombination in compressed SF₆” J. Phys. D: 12. 167-70 1979.
- [III.3] Lemzadmi, AssiaGuerroui, Andre Denat, Nelly Bonifaci: “Mobility of charge carriers in SF₆ and SF₆-N₂ gas mixture at high pressure” Acta Technica 59 (2014), No. 4. pp: 395–404.

Chapitre 4

Réseaux de Neurones

4.1 Introduction

Lorsqu'apparaît une nouvelle technique, l'ingénieur se demande naturellement en quoi cette nouveauté peut lui être utile. Si elle est dotée d'un nom plus métaphorique que technique, (ce qui est évidemment le cas pour les réseaux de neurones), la réponse à cette question doit être particulièrement précise et motivée. De plus, la mise en œuvre des réseaux de neurones est très simple ; la tentation peut être grande, d'appliquer cette technique de manière irréfléchie ou inadaptée, ce qui ne peut conduire qu'à des déceptions. C'est pourquoi nous expliquerons ici les principes fondamentaux qui justifient l'intérêt pratique des réseaux de neurones, et nous situerons ces derniers dans la perspective des méthodes classiques de traitement statistique de données ; nous montrerons que la technique des réseaux de neurones formels doit être considérée comme une extension puissante de techniques bien connues des ingénieurs.

4.2 Historique

Les débuts.

- 1890 : W. James, célèbre psychologue américain introduit le concept de mémoire associative, et propose ce qui deviendra une loi de fonctionnement pour l'apprentissage sur les réseaux de neurones connue plus tard sous le nom de loi de Hebb [IV.1].
- 1943 : J. Mc Culloch et W. Pitts laissent leurs noms à une modélisation du neurone biologique (un neurone au comportement binaire). Ceux sont les premiers à montrer que des réseaux de neurones formels simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes (tout au moins au niveau théorique) [IV.2].
- 1949 : D. Hebb, physiologiste américain explique le conditionnement chez l'animal par les

propriétés des neurones eux-mêmes. Ainsi, un conditionnement de type pavlovien tel que, nourrir tous les jours à la même heure un chien, entraîne chez cet animal la sécrétion de salive à cette heure précise même en l'absence de nourriture. La loi de modification des propriétés des connexions entre neurones qu'il propose explique en partie ce type de résultats expérimentaux [IV.3].

Les premiers succès

- 1957 : F. Rosenblatt développe le modèle du Perceptron. Il construit le premier neuro-ordinateur basé sur ce modèle et l'applique au domaine de la reconnaissance de formes. Notons qu'à cet époque les moyens à sa disposition sont limités et c'est une prouesse technologique que de réussir à faire fonctionner correctement cette machine plus de quelques minutes [IV.4- IV.5].

- 1960 : B. Widrow, un automaticien, développe le modèle Adaline(Adaptative Linear Element). Dans sa structure, le modèle ressemble au Perceptron, cependant la loi d'apprentissage est différente [IV.6]. Celle-ci est à l'origine de l'algorithme de rétropropagation de gradient très utilisé aujourd'hui avec les Perceptrons multicouches. Les réseaux de type Adaline restent utilisés de nos jours pour certaines applications particulières.

- 1969 : M. Minsky et S. Papert publient un ouvrage qui met en exergue les limitations théoriques du perceptron. Limitations alors connues, notamment concernant l'impossibilité de traiter par ce modèle des problèmes non linéaires. Ils étendent implicitement ces limitations à tous modèles de réseaux de neurones artificiels. Leur objectif est atteint, il y a abandon financier des recherches dans le domaine (surtout aux U.S.A.), les chercheurs se tournent principalement vers l'Intelligence Artificielle et les systèmes à bases de règles [IV.7].

L'ombre

- 1967 à 1982 : Toutes les recherches ne sont, bien sûr, pas interrompues. Elles se poursuivent, mais déguisées, sous le couvert de divers domaines comme : le traitement adaptatif du signal, la reconnaissance de formes, la modélisation en neurobiologie, etc. De grands noms travaillent durant cette période telle : S. Grosberg, 1976 [IV.8] et T. Kohonen [IV.9].

Le renouveau

- 1982 : J. J. Hopfield est un physicien reconnu à qui l'on doit le renouveau d'intérêt pour les réseaux de neurones artificiels. A cela plusieurs raisons : Au travers d'un article court, clair et bien écrit, il présente une théorie du fonctionnement et des possibilités des réseaux de neurones [IV.10]. Il faut remarquer la présentation anticonformiste de son article. Alors que les auteurs s'acharnent jusqu'alors à proposer une structure et une loi d'apprentissage, puis à étudier les propriétés émergentes ; J. J. Hopfield fixe préalablement le comportement à atteindre pour son modèle et construit à partir de là, la structure et la loi d'apprentissage correspondant au résultat escompté. Ce modèle est aujourd'hui encore très utilisé pour des problèmes d'optimisation. D'autre part, entre les mains de ce physicien distingué, la théorie des réseaux de neurones devient respectable. Elle n'est plus l'apanage d'un certain nombre de psychologues et neurobiologistes hors du coup.

Notons qu'à cette date, l'intelligence artificielle est l'objet d'une certaine désillusion, elle n'a pas répondu à toutes les attentes et s'est même heurtée à de sérieuses limitations. Aussi, bien que les limitations du Perceptron mise en avant par M. Minsky ne soient pas levées par le modèle d'Hopfield, les recherches sont relancées.

La levée des limitations

- 1985 : La Machine de Boltzmann est le premier modèle connu apte à traiter de manière satisfaisante les limitations recensées dans le cas du perceptron. Mais l'utilisation pratique s'avère difficile, la convergence de l'algorithme étant extrêmement longue (les temps de calcul sont considérables) [IV.11].

- 1986 : L'algorithme d'apprentissage de rétro-propagation de gradient apparaît [IV.12]. C'est un algorithme d'apprentissage adapté aux réseaux de neurones multicouches (aussi appelés Perceptrons multicouches). Dès cette découverte, nous avons la possibilité de réaliser une fonction non linéaire d'entrée/sortie sur un réseau en décomposant cette fonction en une suite d'étapes linéairement séparables. De nos jours, les réseaux multicouches et la rétro-propagation de gradient reste le modèle le plus étudié et le plus productif au niveau des applications.-1989 : Vingt ans après la publication de l'ouvrage où Minsky et Papert exposaient les limitations du perceptron, [IV.13- IV.14 - IV.15] établissent les réseaux de neurones comme une classe

d'approximateurs universels. Il a été ainsi démontré qu'un perceptron multicouches avec une seule couche cachée pourvue d'un nombre suffisant de neurones, peut approximer n'importe quelle fonction avec la précision souhaitée. Néanmoins, cette propriété ne permet pas de choisir, pour un type de fonction donné, le nombre de neurones optimal dans la couche cachée. En d'autres termes, ce résultat ne mène pas vers une technique de construction d'architecture.

-1994 : Hornik [IV.16] a démontré la parcimonie des réseaux de neurones à une couche cachée, ce qui a permis d'étendre le champ d'application des réseaux de neurones aux applications en temps réel [IV.17]

La situation actuelle

Les réseaux de neurones occupent de nos jours une place importante dans le domaine de la modélisation, de la commande et de la classification, ce qui justifie l'intérêt que le monde scientifique et industriel continue à porter au « connexionnisme ».

4.3 Les neurones formels

Un "neurone formel" (ou simplement "neurone") est une fonction algébrique non linéaire et bornée, dont la valeur dépend de paramètres appelés coefficients ou poids. Les variables de cette fonction sont habituellement appelées "entrées" du neurone, et la valeur de la fonction est appelée sa "sortie".

Un neurone est donc avant tout un opérateur mathématique, dont on peut calculer la valeur numérique par quelques lignes de logiciel. On a pris l'habitude de représenter graphiquement un neurone comme indiqué sur la Figure 4.1.

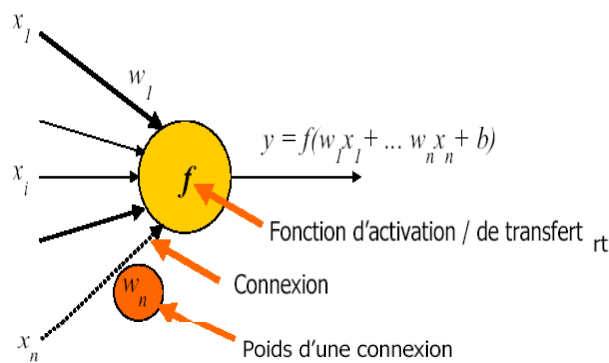


Figure 4.1: Neurone formel

Où les x_i sont les variables (ou entrées) du neurone, les w_i sont des paramètres ajustables, y est la sortie du neurone et b est le *biais*.

Les neurones les plus fréquemment utilisés sont ceux pour lesquels la fonction d'activation f est une fonction non linéaire (généralement une *tangente hyperbolique* ou une *sigmoïde*) d'une combinaison linéaire des entrées.

4.3.1 Différents types de neurones

La fonction d'activation (de transfert) utilisée dans le modèle de McCulloch&Pitts est la fonction *échelon* (Figure 4.2-a). Elle fait passer l'activation du neurone d'une valeur à une autre dès que l'entrée résultante dépasse un certain seuil. L'inconvénient de cette fonction est qu'elle n'est pas différentiable, ce qui pose un problème pour les algorithmes basés sur le gradient.

Pour remédier à cet inconvénient, on cherche à approximer cette fonction d'activation par une fonction « non linéaire » différentiable.

Deux fonctions de ce type sont particulièrement intéressantes et sont souvent utilisées : la fonction tangente hyperbolique (Figure 4.2.b) définie par :

$$f(u) = \tanh(\beta u) = \frac{e^{\beta u} - e^{-\beta u}}{e^{\beta u} + e^{-\beta u}} \quad (4.1)$$

et la fonction logistique (Figure 4.2.c) dont l'expression est la suivante :

$$f_{\beta}(u) = \frac{1}{1 + e^{-\beta u}} \quad (4.2)$$

La fonction « tanh » est bornée entre -1 et +1 alors que la fonction logistique est bornée entre 0 et 1. Ces deux fonctions, appelées *fonctions sigmoïdes*, sont liées par la relation :

$$\tanh(\beta u) = 2f_{\beta}(u) - 1 \quad (4.3)$$

où le paramètre β est appelé le gain. Plus le gain est important, plus la saturation du neurone est rapide (dans notre travail, le paramètre β est pris égal à 1).

La fonction logistique (4.2) est appelée aussi *sigmoïde unipolaire*. Elle admet une variante appelée *sigmoïde bipolaire* (Figure 4.2.d) ayant pour expression :

$$f_{\beta}(u) = \frac{1 - e^{-\beta u}}{1 + e^{-\beta u}} \quad (4.4)$$

Les fonctions sigmoïdes ont la propriété d'être différentiables, ce qui est nécessaire pour les algorithmes basés sur le gradient. Une autre propriété intéressante est le fait que les fonctions dérivées peuvent s'exprimer facilement à l'aide des fonctions elles-mêmes, ce qui permet un gain significatif de temps de calcul [IV.18]

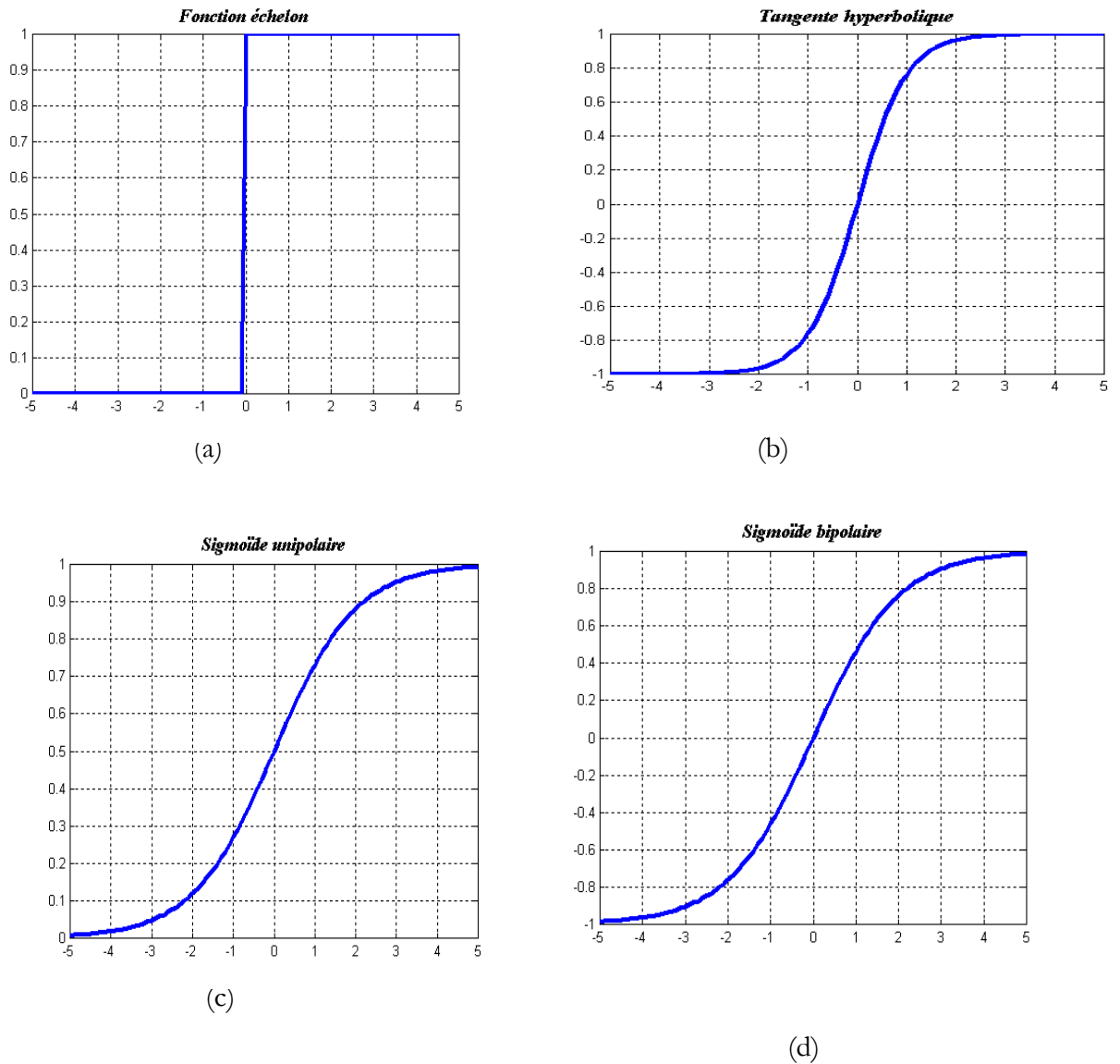


Figure 4.2: Différentes fonctions d'activation

Un neurone formel ne réalise donc rien d'autre qu'une somme pondérée suivie d'un non linéarité. C'est l'association de tels éléments simples sous la forme de réseaux qui permet de réaliser des fonctions utiles pour des applications industrielles.

4.4 Les réseaux de neurones formels

On distingue deux grands types d'architectures de réseaux de neurones : les réseaux de neurones *non bouclés (statiques)* et les réseaux de neurones *bouclés (dynamiques ou récurrents)*.

4.4.1 Les réseaux de neurones non bouclés (statiques)

Un réseau de neurones non bouclé réalise une (ou plusieurs) fonctions algébriques de ses entrées, par composition des fonctions réalisées par chacun de ses neurones.

Un réseau de neurones non bouclé est représenté graphiquement par un ensemble de neurones connectés entre eux, l'information circulant des entrées vers les sorties sans retour en arrière ; si l'on représente le réseau comme un graphe dont les nœuds sont les neurones et les arêtes les connexions entre ceux-ci, le graphe d'un réseau non bouclé est *acyclique*. Le terme de connexions est une métaphore : dans la très grande majorité des applications, les réseaux de neurones sont des formules algébriques dont les valeurs numériques sont calculées par des programmes d'ordinateurs, non des objets physiques (circuits électroniques spécialisés) ; néanmoins, le terme de connexion, issu des origines biologiques des réseaux de neurones, est passé dans l'usage, car il est commode quoique trompeur. Il a même donné naissance au terme de « connexionnisme ».

4.4.1.1 Perceptron multicouches (PMC)

La Figure 4.3 représente un réseau de neurones non bouclé qui a une structure particulière, très fréquemment utilisée : c'est le PMC. Ses neurones sont organisés en couches successives. Chaque neurone d'une couche reçoit des signaux de la couche précédente et transmet le résultat à la suivante, si elle existe.

Les neurones d'une même couche ne sont pas interconnectés. Un neurone ne peut donc envoyer son résultat qu'à un neurone situé dans une couche postérieure à la sienne.

L'orientation du réseau est fixée par le sens, unique, de propagation de l'information, de la couche d'entrée vers la couche de sortie.

Pour le réseau considéré, les notions de couches d'entrée et de sortie sont donc systématiques. Ces dernières constituent l'interface du réseau avec l'extérieur. La couche d'entrée reçoit les signaux (ou variables) d'entrée et la couche de sortie fournit les résultats. Enfin, les neurones des autres couches (couches cachées) n'ont aucun lien avec l'extérieur et sont appelés neurones cachés.

Par convention, les neurones d'entrée ont toujours une fonction d'activation « identité », laissant passer l'information sans la modifier.

En ce qui concerne les neurones de sortie, on peut leur associer une fonction d'activation linéaire ou non, dérivable ou non, suivant la nature du problème à résoudre.

En ce qui concerne la fonction d'activation associée aux neurones cachés, on utilise dans le cadre de cette thèse une fonction d'activation de la famille des sigmoïdes.

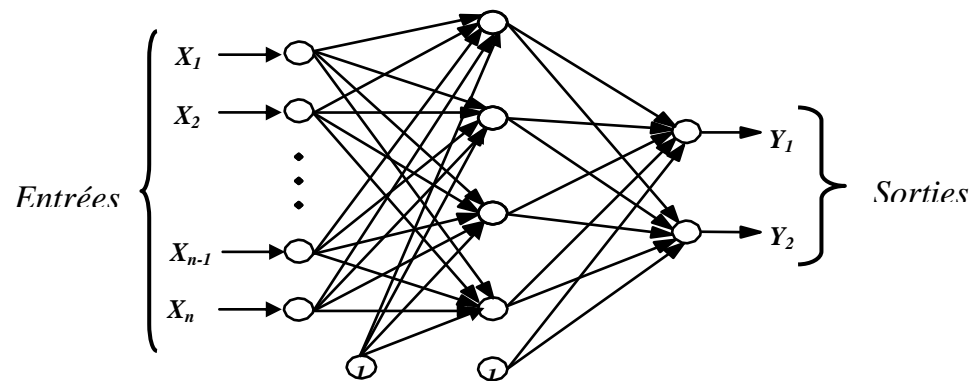


Figure 4.3: Perceptron multicouche à une couche cachée

Les réseaux de neurones non bouclés sont des objets statiques : si les entrées sont indépendantes du temps, les sorties le sont également. Ils sont utilisés principalement pour effectuer des tâches d'approximation de fonction non linéaire, de classification ou de modélisation de processus statiques non linéaires.

4.4.2 Les réseaux de neurones bouclés (dynamiques)

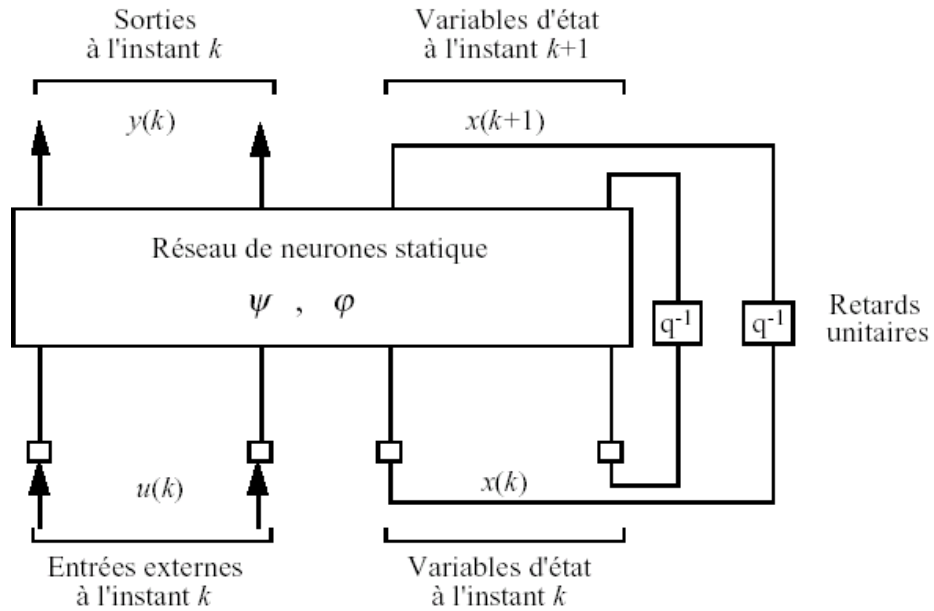
Contrairement aux réseaux de neurones non bouclés dont le graphe de connexions est acyclique, les réseaux de neurones bouclés peuvent avoir une topologie de connexions quelconque, comprenant notamment des boucles qui ramènent aux entrées la valeur d'une ou plusieurs sorties. Pour qu'un tel système soit causal, il faut évidemment qu'à toute boucle soit associé un retard : un réseau de neurones bouclé est donc un système dynamique, régi par des équations différentielles ; comme l'immense majorité des applications sont réalisées par des programmes d'ordinateurs, on se place dans le cadre des systèmes à temps discret, où les équations différentielles sont remplacées par des équations aux différences.

Un réseau de neurones bouclé à temps discret est donc régi par une (ou plusieurs) équations aux différences non linéaires, résultant de la composition des fonctions réalisées par chacun des neurones et des retards associés à chacune des connexions.

La forme la plus générale des équations régissant un réseau de neurones bouclé est appelée forme canonique ;

$$\begin{aligned}x(k+1) &= \varphi(x(k), u(k)) \\ y(k) &= \psi(x(k), u(k))\end{aligned}\tag{4.5}$$

Où φ et ψ sont des fonctions non linéaires réalisées par un réseau de neurones non bouclé (par exemple, mais pas obligatoirement, un Perceptron multicouche), et où k désigne le temps (discret). La forme canonique est représentée sur la Figure 4.4. Tout réseau de neurones, bouclé aussi compliqué soit-il, peut être mis sous cette forme canonique [IV.19]



L'état du réseau de neurones bouclé peut être exprimé en fonction des valeurs passées de la sortie.

Les réseaux de neurones bouclés sont utilisés pour effectuer des tâches de modélisation de systèmes dynamiques, de commande de processus, ou de filtrage.

4.4.3 Propriétés des Réseaux de Neurones Formels

Les propriétés essentielles des réseaux de neurones sont :

4.4.3.1 Le parallélisme

Cette notion se situe à la base de l'architecture des réseaux de neurones considérés comme ensemble d'entités élémentaires qui travaillent simultanément.

Le parallélisme permet une rapidité de calcul supérieure mais exige de penser et de poser différemment les problèmes à résoudre.

4.4.3.2 **La capacité d'adaptation**

Elle se manifeste par la capacité d'apprentissage qui permet de tenir compte des nouvelles contraintes ou de nouvelles données du monde extérieur. Certains réseaux se caractérisent aussi par leur capacité d'auto-organisation qui assure leur stabilité en tant que systèmes dynamiques capables de tenir compte de situations non encore connues.

4.4.3.3 **La facilité de construction**

La simulation informatique de réseaux de neurones pour une petite application est simple et ne nécessite qu'un temps de développement assez court. Pour des applications plus complexes, l'utilisation de simulateur ou de carte accélératrice se révèle utile.

4.4.3.4 **Propriété fondamentale : L'Approximation universelle parcimonieuse**

4.4.3.4.1 **L'approximation universelle**

Les travaux de Cybenko (Cybenko,1989) et Funahashi (Funahashi, 1989) [IV.14 - IV.15] ont prouvé la possibilité d'approcher des fonctions continues, au sens de la norme uniforme sur les compacts, par des réseaux de neurones.

Les réseaux considérés sont de type réseau à une couche de neurones cachés à fonction d'activation non linéaire, et à neurones de sortie linéaires. Dans le cas du théorème de Cybenko, l'hypothèse sur la fonction d'activation est qu'elle a pour limite 0 en $-\infty$ et 1 en $+\infty$, dans celui de Funahashi, qu'elle est croissante, non constante et bornée. Les fonctions non continues, mais mesurables, peuvent aussi être approchées, mais dans un sens moins fort [IV.20]. Il existe par ailleurs quelques résultats sur le nombre de neurones requis pour approcher une fonction avec une précision fixée, pour certaines classes de fonctions [IV.21].

Ces résultats affirment donc, pour toute fonction déterministe usuelle, l'existence d'une approximation par un réseau de neurones. Les réseaux complètement connectés ou à couches, et à neurones cachés sigmoïdaux, remplissent les conditions d'application des théorèmes.

Dans ce travail, nous utilisons systématiquement ce type de réseaux, à l'exclusion par exemple des réseaux utilisant des fonctions à base radiale (RBF). Une raison en est que, même si ces réseaux jouissent également de propriétés d'approximation intéressantes, et même si leur

Apprentissage peut être réalisé à l'aide de la méthode des moindres carrés ordinaires, leur utilisation est souvent beaucoup moins économique, ou parcimonieuse du point de vue du nombre de connexions, que celle des réseaux à sigmoïdes. En toute rigueur, les réseaux à RBF peuvent être aussi parcimonieux que les réseaux à sigmoïdes, mais à condition d'ajuster la position des centres des RBF de manière non linéaire [IV.22], ce qui supprime le principal intérêt des RBF : la simplicité de l'apprentissage par la méthode des moindres carrés ordinaires.

4.4.3.4.2 La parcimonie

Lorsque l'on cherche à modéliser un processus à partir des données, on s'efforce toujours d'obtenir les résultats les plus satisfaisants possibles avec un nombre minimum de paramètres ajustables. Dans cette optique, [IV.16].a montré que : *Si le résultat de l'approximation (c'est-à-dire la sortie du réseau de neurones) est une fonction non linéaire des paramètres ajustables, elle est plus parcimonieuse que si elle est une fonction linéaire de ces paramètres. De plus, pour des réseaux de neurones à fonction d'activation sigmoïdale, l'erreur commise dans l'approximation varie comme l'inverse du nombre de neurones cachés, et elle est indépendante du nombre de variables de la fonction à approcher. Par conséquent, pour une précision donnée, donc pour un nombre de neurones cachés donné, le nombre de paramètres du réseau est proportionnel au nombre de variables de la fonction à approcher.*

Ce résultat s'applique aux réseaux de neurones à fonction d'activation sigmoïdale puisque la sortie de ces neurones n'est pas linéaire par rapports aux poids synaptiques.

La spécificité des réseaux de neurones réside donc dans le caractère parcimonieux de l'approximation : à précision égale, les réseaux de neurones nécessitent moins de paramètres ajustables (les poids des connexions) que les approximateurs universels couramment utilisés ; plus précisément, le nombre de poids varie *linéairement* avec le nombre de variables de la fonction à approcher, alors qu'il varie *exponentiellement* pour la plupart des autres approximateurs [IV.21].

Qualitativement, la propriété de parcimonie peut se comprendre de la manière suivante : lorsque l'approximation est une combinaison *linéaire* de fonctions élémentaires fixées (des monômes par exemple, ou des gaussiennes à centres et écarts-types fixes), on ne peut ajuster que les coefficients de la combinaison ; en revanche, lorsque l'approximation est une combinaison linéaire de fonctions non linéaires à paramètres ajustables (un perceptron multicouche par exemple), on ajuste à la fois les coefficients de la combinaison et la forme des fonctions que l'on combine. Ainsi, dans un perceptron multicouche, les poids de la première

Couche déterminent la forme de chacune des sigmoïdes réalisées par les neurones cachés, et les poids de la seconde couche déterminent une combinaison linéaire de ces fonctions. On conçoit facilement que cette souplesse supplémentaire, conférée par le fait que l'on ajuste la forme des fonctions que l'on superpose, permet d'utiliser un plus petit nombre de fonctions élémentaires, donc un plus petit nombre de paramètres ajustables. Nous allons voir ultérieurement pourquoi cette propriété de parcimonie est précieuse dans les applications industrielles.

Rappelons que ces résultats concernent l'utilisation de réseaux de neurones pour l'approximation uniforme de fonctions connues ; il est pourtant rare que les réseaux de neurones soient mis en œuvre dans ce cadre. Nous allons montrer dans le paragraphe suivant que la technique des réseaux de neurones est généralement utilisée comme une méthode de modélisation statistique.

4.5 Réseaux de neurones et régressions non linéaires

Dans la pratique, on n'utilise pas les réseaux de neurones pour réaliser des approximations de fonctions connues. Le plus souvent, le problème qui se pose à l'ingénieur est le suivant : il dispose d'un ensemble de mesures de variables d'un processus de nature quelconque (physique, chimique, économique, financier, ...), et du résultat de ce processus ; il suppose qu'il existe une relation déterministe entre ces variables et ce résultat, et il cherche une forme mathématique de cette relation, valable dans le domaine où les mesures ont été effectuées, sachant que les mesures sont en nombre fini, qu'elles sont certainement entachées de bruit, et que toutes les variables qui déterminent le résultat du processus ne sont pas forcément mesurées.

En d'autres termes, l'ingénieur cherche un modèle du processus qu'il étudie, à partir des mesures dont il dispose, et d'elles seules : on dit qu'il effectue une modélisation « boîte noire ».

Dans le jargon des réseaux de neurones, les données à partir desquelles on cherche à construire le modèle s'appellent *des exemples*.

En quoi la propriété d'approximation parcimonieuse peut-elle être utile pour résoudre ce genre de problèmes ? Ce que l'ingénieur cherche à obtenir à l'aide de son modèle, c'est la vraie fonction qui relie la grandeur de sortie que l'on veut modéliser aux variables d'entrées qui la déterminent, en d'autres termes, la fonction que l'on obtiendrait en faisant une infinité de mesures de la grandeur de sortie pour chaque valeur possible de variables d'entrées : en termes de statistiques, l'ingénieur cherche la fonction de régression de la grandeur à modéliser. Cette

fonction est inconnue, mais on peut en chercher une approximation à partir des mesures disponibles : les réseaux de neurones sont donc de bons candidats pour cela, si la fonction de régression cherchée est non linéaire. Cette approximation est obtenue en estimant les paramètres d'un réseau de neurones au cours d'une phase dite *d'apprentissage*. C'est ici que la propriété d'approximation parcimonieuse des réseaux de neurones est précieuse : en effet, le nombre de mesures nécessaires pour estimer les paramètres de manière significative est d'autant plus grand que le nombre de paramètres est grand. Ainsi, pour modéliser une grandeur avec une précision donnée à l'aide d'un réseau de neurones, il faut *moins de données* que pour la modéliser, avec une précision comparable, à l'aide d'une régression linéaire multiple ; de manière équivalente, un réseau de neurones permet, avec les mêmes données disponibles, de réaliser une approximation plus précise qu'une régression linéaire multiple [IV.17]. De manière générale, un réseau de neurones permet donc de faire un meilleur usage des mesures disponibles que les méthodes de régression non linéaires conventionnelles. Ce gain peut être considérable lorsque le processus à modéliser dépend de plusieurs variables.

Ainsi, à la lumière de cette propriété fondamentale, la technique des réseaux de neurones apparaît comme une puissante méthode de régression non linéaire : ce n'est donc rien d'autre qu'une extension des méthodes de régression linéaire ou multilinéaires proposées par tous les logiciels qui permettent de faire de la modélisation de données.

Contrairement à une croyance répandue, la technique des réseaux de neurones ne relève donc pas de l'Intelligence Artificielle au sens classique du terme, mais elle constitue une branche des statistiques appliquées. Il ne faut donc pas être victime du vocabulaire utilisé (neurones, apprentissage, etc.) ; Le tableau ci-dessous résume les équivalences entre le vocabulaire des statistiques et celui des réseaux de neurones.

RESEAUX DE NEURONES	STATISTIQUES
Choix de l'architecture	Choix de la famille de fonctions destinées à approcher la fonction de régression
Ensemble d'apprentissage	Observations
Apprentissage	Estimation des paramètres de l'approximation de la fonction de régression
Généralisation	Interpolation, extrapolation

Tableau 4.1 : Réseaux de neurones et statistiques

4.6 Apprentissage d'un réseau de neurones

Un réseau de neurones définit une famille de fonctions. L'apprentissage consiste à déterminer la solution du problème posé au sein de cette famille de fonctions. Ces fonctions pourront avoir des capacités limitées comme les fonctions linéaires ou au contraire permettre la construction de fonctions aussi complexes qu'on le désire comme les PMCs. Le principe d'apprentissage est l'optimisation d'une *fonction de coût* calculée à partir des exemples de la base d'apprentissage et de la sortie du réseau de neurones.

Les méthodes numériques utilisées sont le plus souvent des méthodes approchées basées sur des techniques de gradient.

En d'autres termes, l'apprentissage d'un réseau de neurones consiste à déterminer les valeurs des poids permettant à la sortie du réseau de neurones d'être aussi proche que possible de l'objectif fixé.

4.6.1 Type d'apprentissage

Il existe de nombreux types de règles d'apprentissage qui peuvent être regroupées en trois catégories: les règles d'apprentissage supervisé, non supervisé, et renforcé. Mais l'objectif fondamental de l'apprentissage reste le même : soit la classification, l'approximation de fonction ou encore la prévision. Dans l'optique de la prévision, l'apprentissage consiste à extraire des régularités (à partir des exemples) qui peuvent être transférées à de nouveaux exemples.

✓ *Apprentissage supervisé*

Un apprentissage est dit supervisé lorsque l'on force le réseau à converger vers un état final précis, en même temps qu'on lui présente un motif. Ce genre d'apprentissage est réalisé à l'aide d'une base d'apprentissage, constituée de plusieurs exemples de type entrées-sorties (les entrées du réseau et les sorties désirées ou encore les solutions souhaitées pour l'ensemble des sorties du réseau).

La procédure usuelle dans le cadre de la prévision est l'apprentissage supervisé (ou à partir d'exemples) qui consiste à associer une réponse spécifique désirée à chaque signal d'entrée. La modification des poids s'effectue progressivement jusqu'à ce que l'erreur (ou l'écart) entre les sorties du réseau (ou résultats calculés) et les résultats désirés soient minimisés.

Cet apprentissage n'est possible que si un large jeu de données est disponible et si les solutions sont connues pour les exemples de la base d'apprentissage.

✓ *Apprentissage renforcé*

L'apprentissage renforcé est une technique similaire à l'apprentissage supervisé à la différence qu'au lieu de fournir des résultats désirés au réseau, on lui accorde plutôt un grade (ou score) qui est une mesure du degré de performance du réseau après quelques itérations. Les algorithmes utilisant la procédure d'apprentissage renforcé sont surtout utilisés dans le domaine des systèmes de contrôle [IV.23].

✓ *Apprentissage non supervisé*

L'apprentissage non supervisé consiste à ajuster les poids à partir d'un seul ensemble d'apprentissage formé uniquement de données. Aucun résultat désiré n'est fourni au réseau.

Qu'est-ce que le réseau apprend exactement dans ce cas ? L'apprentissage consiste à

Détecter les similarités et les différences dans l'ensemble d'apprentissage. Les poids et les sorties du réseau convergent, en théorie, vers les représentations qui capturent les régularités statistiques des données. Ce type d'apprentissage est également dit compétitif et (ou) coopératif (Grossberg, 1988). L'avantage de ce type d'apprentissage réside dans sa grande capacité d'adaptation reconnue comme une auto organisation, « self-organizing » [IV.24]. L'apprentissage non supervisé est surtout utilisé pour le traitement du signal et l'analyse factorielle.

4.6.2 Algorithme de Rétro-propagation (RP)

L'algorithme de *rétropropagation du gradient* (Back-propagation, en Anglais) est certainement à la base des premiers succès des réseaux de neurones. Sa mise en application a permis au domaine du « connexionnisme » de sortir de la période de silence qui a régné après la sortie du livre « Perceptrons » de [IV.7]. Il figure aujourd'hui parmi les algorithmes d'apprentissage les plus utilisés.

Selon la littérature, la *RP* a été proposée plusieurs fois et de manière indépendante par : ([IV.26-IV.27]), et enfin Rumelhart et les membres du groupe PDP en 1986 [IV.12]. Une approche similaire a également été proposée par [IV.28].

D'après [IV.29], des liens peuvent être établis avec la technique de [IV.30]. Cependant, la popularisation de la *RP* et son développement reste lié aux travaux du groupe PDP dirigé par Rumelhart.

On considère un réseau à trois couches illustré par la Figure 4.5.

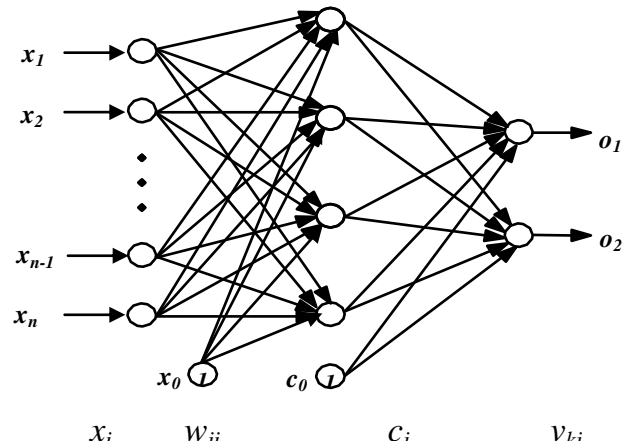


Figure 4.5: Définition des notations pour un PMC

Les conventions de notation sont les suivantes :

- o_k : activation de la $k^{\text{ème}}$ unité de sortie, $k = 1, \dots, M$;
- k_T : activation désirée de la $k^{\text{ème}}$ unité de sortie;
- c_j : activation de la $j^{\text{ème}}$ unité cachée, $j = 0, 1, \dots, n$; $c_0 = 1$: c'est l'entrée du biais pour la couche de sortie;
- x_i : $i^{\text{ème}}$ entrée externe du réseau ; $i = 0, 1, \dots, n$; $x_0 = 1$: entrée du biais pour la couche cachée;
- w_{ji} : poids d'une connexion entre la $i^{\text{ème}}$ entrée et la $j^{\text{ème}}$ unité cachée;
- v_{kj} : poids d'une connexion entre la $j^{\text{ème}}$ unité cachée et la $k^{\text{ème}}$ unité de sortie.

Les indices i, j et k font référence aux unités d'entrée, aux unités cachées et aux unités de sortie, respectivement. L'exposant p correspond au numéro de l'exemple présenté à l'entrée du réseau : $p = 1, \dots, n_A$, où n_A est le nombre d'exemples d'apprentissage. Le $p^{\text{ème}}$ exemple est noté $x^p = x_0^p, \dots, x_i^p, \dots, x_n^p$ et la $i^{\text{ème}}$ composante x_i^p désigne la $i^{\text{ème}}$ entrée lorsque le $p^{\text{ème}}$ exemple est présenté au réseau. Pour un exemple p , la $j^{\text{ème}}$ unité cachée a l'entrée résultante I_j^p :

$$I_j^p = \sum_{i=0}^n w_{ji} x_i^p \quad (4.6)$$

et une activation c_j^p :

$$c_j^p = f(I_j^p) = f\left(\sum_{i=0}^n w_{ji} x_i^p\right) \quad (4.7)$$

où f est la fonction d'activation. La $k^{\text{ème}}$ unité de sortie reçoit une entrée résultante I_k^p définie par :

$$I_k^p = \sum_{j=0}^n v_{kj} c_j^p \quad (4.8)$$

et génère en sortie l'activation o_k^p :

$$o_k^p = f(I_k^p) \quad (4.9)$$

Comme la fonction d'activation des neurones de sorties pour un PMC est linéaire, l'équation (4.9) devient :

$$o_k^p = I_k^p \quad (4.10)$$

La fonction de coût usuelle est l'erreur quadratique moyenne définie par :

$$E(\mathbf{w}) = \frac{1}{2} \sum_{k,p} (t_k^p - o_k^p)^2 \quad (4.11)$$

où \mathbf{w} est le vecteur contenant tous les poids du réseau. La fonction $E(\mathbf{w})$ est continue et différentiable par rapport à chaque poids. Pour déterminer les poids qui la minimisent, on peut donc utiliser l'algorithme de descente du gradient. Pour faciliter la notation, $E(\mathbf{w})$ sera notée E dans ce qui suit [IV.31].

Pour les poids des connexions des unités cachées vers les unités de sortie, le terme d'adaptation des poids au cours de l'apprentissage est défini par :

$$\begin{aligned}\Delta v_{kj} &= -\eta \frac{\partial E}{\partial v_{kj}} \\ &= \eta \sum_p \delta_k^p c_j^p\end{aligned}\quad (4.12)$$

avec :

$$\delta_k^p = (t_k^p - o_k^p) \quad (4.13)$$

dans le cas de sorties linéaires.

Pour les poids des connexions entre la couche d'entrée et la couche cachée, le terme d'adaptation des poids est :

$$\begin{aligned}\Delta w_{ji} &= -\eta \frac{\partial E}{\partial w_{ji}} = -\eta \sum_p \frac{\partial E}{\partial c_j^p} \frac{\partial c_j^p}{\partial w_{ji}} \\ &= \eta \sum_{k,p} (t_k^p - o_k^p) v_{kj} f'(I_j^p) x_i^p = \eta \sum_p \delta_j^p x_i^p\end{aligned}\quad (4.14)$$

avec :

$$\delta_j^p = f'(I_j^p) \sum_k v_{kj} \delta_k^p \quad (4.15)$$

On peut constater que les équations (4.12) et (4.14) ont la même forme et ne diffèrent que par la définition de la quantité δ . Ces formules se généralisent facilement aux cas des réseaux possédant un nombre quelconque de couches cachées.

D'après l'équation (4.15), le calcul de δ_j pour une unité cachée j nécessite les δ_k des unités de sortie, qui sont fonctions des erreurs en sortie du réseau ($t_k - o_k$). Ainsi, pour corriger les poids des connexions entre la couche d'entrée et la couche cachée, on a besoin de rétropropager

L'erreur depuis les sorties vers les entrées, d'où le nom de l'algorithme d'apprentissage : *rétro-propagation de l'erreur*.

D'après l'équation (4.15), le calcul de δ_j pour une unité cachée j utilise la dérivée de l'activation de cette même unité et la somme des δ_k pour toutes les unités de sortie. Le calcul de δ_j se fait donc de manière locale (indépendamment de toutes les autres unités cachées). Ceci permet d'envisager une *parallélisations* de l'algorithme de rétro-propagation (RP).

Après avoir calculé le terme d'adaptation des poids, la mise à jour se fait selon la formule suivante :

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji} \quad (4.16)$$

Et

$$v_{kj}(t+1) = v_{kj}(t) + \Delta v_{kj} \quad (4.17)$$

Où t est l'indice de l'itération. Le mode d'adaptation des poids tel qu'il est présenté par les équations (4.12) et (4.14) s'appelle mode « *batch* ». La mise à jour des poids se fait après avoir passée en revue tous les exemples d'apprentissage. Ce mode d'apprentissage est encore appelé déterministe hors ligne « *off-line* » ou « *by epoch* ».

Une autre approche consiste à modifier les poids après chaque présentation d'un exemple d'apprentissage. C'est l'apprentissage en mode en ligne « *on-line* » ou « *by pattern* ».

Il y a des avantages et des inconvénients pour chaque mode d'adaptation des poids, et le choix de l'un ou de l'autre dépendent du problème à traiter.

Les algorithmes « *off-line* » sont faciles à analyser pour ce qui concerne les propriétés de convergence. Ils peuvent utiliser un taux d'apprentissage optimum à chaque itération et peuvent conduire à des solutions assez précises (avec de faibles variantes). En revanche, ils ont l'inconvénient d'induire un temps de calcul du terme d'adaptation de poids dépendant de la taille de l'ensemble d'apprentissage.

Les méthodes « *on-line* » peuvent être utilisées lorsque les exemples ne sont pas tous disponibles au début de l'apprentissage, et quand on désire réaliser une adaptation continue à partir d'une suite de couples *entrée-sortie* issus d'une relation qu'on cherche à identifier.

4.6.3 Variantes de l'algorithme de RP

Depuis son introduction, l'algorithme de RP a été largement étudié et plusieurs modifications y ont été apportées. L'algorithme de base décrit ci-dessus converge très lentement pour les réseaux multicouches.

Les variations apportées à l'algorithme de RP ont pour objectifs l'accélération de la convergence du processus d'apprentissage et l'amélioration de la capacité de généralisation. Nous présentons ci-dessous quelques variantes parmi les plus importantes.

4.6.3.1 Choix de la fonction coût

Le choix de la fonction d'erreur utilisée pour l'apprentissage des réseaux de neurones multicouches a une certaine influence sur la rapidité d'apprentissage et sur la qualité de généralisation du réseau. Cette question a été étudiée par plusieurs chercheurs [IV.32]. Le critère d'erreur le plus utilisé est la fonction d'erreur quadratique moyenne (cf. équation (4.11).

Cette fonction a tendance à amplifier les erreurs les plus importantes. Par conséquent, au cours de l'apprentissage, la mise à jour des poids est largement déterminée par la correction des grandes erreurs, ce qui est recherché en général [IV.29].

Cependant le choix de fonction quadratique n'est pas la seule possibilité. On peut remplacer $(t_k^p - o_k^p)^2$ par tout autre fonction $\mathcal{E}(t_k^p, o_k^p)$ différentiable et minimale lorsque ses deux arguments sont égaux. Le développement précédent montre que seule l'expression de l'équation (4.12) dépend de la fonction coût. Le reste de l'algorithme de RP reste inchangé.

4.6.3.2 Introduction d'un terme de moment

Le paramètre η (appelé taux d'apprentissage) joue un rôle important. S'il est trop faible, la convergence est lente, et, s'il est trop grand, l'algorithme oscille entre des points différents à cause de l'existence de vallées et de plateaux à la surface de la fonction coût.

Pour stabiliser la recherche des poids optimisant la fonction coût, une méthode consiste à ajouter un terme dit de « *moment* » à l'expression d'adaptation des poids [IV.33]. L'idée est de donner une certaine « *inertie* » pour chaque poids, de sorte que sa mise à jour ne se fasse pas de manière brutale.

Ceci permet alors d'utiliser un taux d'apprentissage relativement important sans pour autant augmenter les oscillations de la trajectoire sur la surface d'erreur.

La nouvelle formule d'adaptation des poids est définie par :

$$\Delta w(t+1) = -\eta \frac{\partial E}{\partial w} + \alpha \Delta w(t) \quad (4.18)$$

Où α est le terme de moment dont la valeur est souvent prise proche de 1 (≈ 0.9). Cette méthode peut être utilisée en modes « off-line » ou « on-line ».

4.6.3.3 Taux d'apprentissage adaptatifs

Les paramètres η et α de l'algorithme RP ne sont pas faciles à fixer a priori pour un problème donné. En outre, des valeurs données peuvent être bien adaptées au début de l'apprentissage, sans l'être nécessairement au milieu ou à la fin du processus. Pour résoudre ce problème, certains auteurs ont proposé d'ajuster automatiquement ces paramètres au cours de l'apprentissage.

Plusieurs heuristiques d'adaptation dynamique du taux d'apprentissage ou du terme de moment ont été proposées [IV.34-IV.35-IV.36]. On se contente ici de rappeler le principe des méthodes adaptatives et de citer l'algorithme de Silva et Almeida comme exemple.

Le principe des méthodes adaptatives est d'étudier à chaque itération l'effet de la mise à jour des poids des connexions sur la fonction coût. Si cette fonction augmente, alors le taux d'apprentissage η_{pq} associé au poids w_{pq} doit être réduit. En revanche, si la fonction coût décroît régulièrement suite aux modifications des poids, alors on peut augmenter η_{pq} pour accélérer la recherche de l'optimum. Différents critères peuvent être retenus, comme par exemple la diminution de la fonction coût pendant plusieurs itérations successives. En partant de ce principe, Silva et Almeida ont proposé la règle suivante :

$$\eta_{pq}(t) = \begin{cases} \gamma \eta_{pq}(t-1), & \text{si } \frac{\partial E}{\partial w_{pq}(t)} \frac{\partial E}{\partial w_{pq}(t-1)} > 0 \\ \zeta \eta_{pq}(t-1) & \text{sinon} \end{cases} \quad (4.19)$$

où w_{pq} est un poids quelconque du réseau, γ et ζ sont des paramètres respectivement supérieur et inférieur à 1 (par exemple: $\gamma = 1.2$ et $\zeta = 0.7$). Les auteurs montrent expérimentalement que les valeurs précises de ces deux paramètres ne sont pas décisives. Ils proposent également d'utiliser un terme de moment identique et non adaptatif pour tous les poids du réseau. La procédure d'adaptation des poids est alors la suivante :

$$\left\{ \begin{array}{l} w_{pq}(t+1) = w_{pq}(t) - \eta_{pq}(t) z_{pq}(t) \\ \text{et} \\ z_{pq}(t) = \frac{\partial E}{\partial w_{pq}(t)} + \alpha z_{pq}(t-1) \end{array} \right. \quad (4.20)$$

Où α est le terme de moment.

4.6.3.4 Autres procédures d'optimisation

L'algorithme de *RP* dans sa forme de base utilise la technique de descente du gradient. Celle-ci est parmi les plus simples, mais elle n'est pas très efficace car elle utilise peu d'information sur la surface de l'erreur. Dans la littérature, on trouve une grande quantité de techniques plus sophistiquées [IV.37]. Une synthèse des principales méthodes est présentée dans [IV.38].

En effectuant le développement au second ordre de la fonction coût $E(\mathbf{w})$ autour du point courant \mathbf{w}_0 et en négligeant les termes d'ordre supérieur, on peut écrire :

$$E(\mathbf{w}) = E(\mathbf{w}_0) + (\mathbf{w} - \mathbf{w}_0)^T \nabla E(\mathbf{w}_0) + \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^T \mathbf{H} (\mathbf{w} - \mathbf{w}_0) \quad (4.21)$$

Avec \mathbf{H} la matrice Hessienne calculée au point \mathbf{w}_0 . En dérivant l'équation (4.21), on obtient :

$$\nabla E(\mathbf{w}) = \nabla E(\mathbf{w}_0) + \mathbf{H} (\mathbf{w} - \mathbf{w}_0) \quad (4.22)$$

En annulant la dérivée, l'équation (4.22) donne une estimation de la localisation du minimum :

$$\mathbf{w} = \mathbf{w}_0 - \mathbf{H}^{-1} \nabla E(\mathbf{w}_0) \quad (4.23)$$

En posant $\mathbf{w}_0 = \mathbf{w}(t)$ et $\mathbf{w} = \mathbf{w}(t+1)$, on obtient une procédure itérative d'estimation du vecteur des poids minimisant la fonction coût :

$$\mathbf{w}(t+1) = \mathbf{w}(t) - \mathbf{H}^{-1} \nabla E(\mathbf{w}(t)) \quad (4.24)$$

Cette méthode est appelée *méthode de Newton*. Elle permet une convergence rapide chaque fois que les poids se trouvent au voisinage de la solution. Son inconvénient est la nécessité de calculer \mathbf{H}^{-1} .

Deux solutions ont été proposées face à ce problème. L'une consiste à négliger les termes non diagonaux de \mathbf{H} (*méthode pseudo-Newton*) [IV.39], l'autre à estimer itérativement la matrice \mathbf{H}^{-1} (*méthode quasi-Newton*).

Les méthodes de minimisation présentées jusqu'ici se basent sur l'utilisation de dérivées de la fonction coût. Une autre approche consiste à utiliser des procédures de *minimisation stochastique* [IV.40] ou les *algorithmes génétiques*. Ces algorithmes effectuent une recherche globale et présentent par conséquent moins de risques d'aboutir à des minima locaux assez superficiels. Cependant, ils nécessitent des calculs plus lourds.

Algorithme de Levenberg-Marquardt

Dans cette méthode de rétro propagation de gradient, pour retrouver les poids qui minimisent l'erreur commise, il est connu en optimisation classique que l'utilisation des informations données par les dérivées d'ordre supérieur peut accélérer le processus de recherche de la solution optimale. De nombreux travaux ont tenté d'étendre cette approche aux réseaux de neurones. On cite la méthode de Newton qui utilise les dérivées de second ordre (la matrice Hessienne). Cependant, il est très difficile et très coûteux de calculer la matrice Hessienne dans le cas des réseaux multicouches. L'algorithme Levenberg-Marquardt [IV.41]. fait partie des méthodes dites Quasi-Newton où le Hessien n'est pas calculé de manière explicite mais approché en utilisant une fonction dépendante du gradient. Cet algorithme est parmi les méthodes les plus performantes [IV.42]. Elle est dite du second ordre parce qu'elle prend en considération la dérivée seconde de la fonction de coût.

4.7 Choix des séquences d'apprentissage

Ce paragraphe donne quelques indications sur le choix des séquences d'apprentissage.

4.7.1 Séquence des entrées de commande

4.7.2 Contraintes sur les entrées de commande :

Elles portent sur l'amplitude et le type de signaux de commande que le processus est susceptible de recevoir pendant son fonctionnement. Les amplitudes maximales sont en général faciles à déterminer, car leur ordre de grandeur correspond aux valeurs de saturation des actionneurs, qui peuvent être estimées physiquement (puissance maximale que peut délivrer un moteur, pression maximale d'un circuit de freinage ou d'un mécanisme hydraulique, etc.).

En ce qui concerne le type de signaux à utiliser, un principe général est que les signaux utilisés pour l'identification doivent être de même nature que ceux qui seront calculés par l'organe de commande pendant l'utilisation du processus. L'idéal serait d'utiliser pour

L'identification le correcteur même qui sera synthétisé à l'aide du prédicteur identifié.

Si le bruit est négligeable, une bonne démarche consiste à effectuer les expériences *en asservissement*, avec un correcteur simple (par exemple, un PID). Cette démarche, ou identification *en boucle fermée*, permet d'explorer le domaine de fonctionnement désiré, en imposant une séquence de consigne correspondant au cahier des charges.

Une autre démarche, couramment utilisée, consiste à explorer au mieux le domaine de fonctionnement, par exemple avec des créneaux de commande (riches en fréquences), d'amplitudes et de durées diverses. Notons que cette dernière solution n'est pas praticable s'il existe des contraintes sur des sorties ou des variables internes du processus, et en particulier si le processus est instable.

4.7.2.1 Fréquence d'échantillonnage :

Les réseaux de neurones étant des modèles non linéaires, il n'est pas possible de passer simplement d'un modèle discret, valable à une fréquence d'échantillonnage donnée, à un autre : il est donc nécessaire d'effectuer l'identification à la fréquence qui sera utilisée pour la commande du processus.

Si des contraintes diverses, par exemple le temps de calcul de la commande, nécessitent le choix d'une fréquence d'asservissement plus basse que la fréquence utilisée pour l'identification, il faut procéder à une nouvelle identification (un nouvel apprentissage).

4.7.3 Séquences d'apprentissage et estimation de la performance

Comme nous l'avons précisé auparavant, l'apprentissage consiste à ajuster les fonctions du prédicteur à un ensemble de points définis par des séquences d'apprentissage, ceci en minimisant une fonction de coût.

Il y a *surajustement* (overfitting) lorsque l'apprentissage conduit à annuler quasiment la fonction de coût sans que pour autant les fonctions réalisées par le réseau prédicteur approchent celles du modèle-hypothèse auquel le prédicteur est associé. Pour des séquences d'apprentissage données, le surajustement se produit si le réseau prédicteur possède trop de coefficients, c'est-à-dire définit une famille de fonctions trop riche.

Pour sélectionner un réseau candidat pour une hypothèse donnée, il est donc nécessaire de répartir les données disponibles en une séquence d'apprentissage et une séquence d'estimation de la performance, dite séquence de test. Une séquence de test, de même type (issue de la même population) que la séquence d'apprentissage, conduit à une meilleure estimation de la variance de l'erreur de prédiction (erreur quadratique moyenne de test, notée $EQMP$) que celle obtenue avec la séquence d'apprentissage (notée $EQMA$).

L'évolution de $EQMA$ et de $EQMP$ lorsque l'on augmente le nombre de neurones cachés, permet de détecter un surajustement, et de sélectionner le meilleur réseau de neurones parmi les candidats : dès qu'ajouter un neurone supplémentaire au réseau fait augmenter l' $EQMP$, même si l' $EQMA$ continue à décroître, le nombre optimal de neurones est atteint.

Cette méthode, à condition d'utiliser des méthodes d'optimisation performantes (par exemple les méthodes quasi-newtoniennes), permet de sélectionner le réseau le plus parcimonieux.

Il est à noter au passage que dans la littérature [IV.43], la séquence de test est souvent utilisée pour arrêter l'apprentissage d'un réseau donné : dès qu'une itération d'apprentissage augmente l' $EQMP$, celui-ci est arrêté prématurément (*Early Stopping*).

De même, la séquence de test peut servir exclusivement à estimer la performance des réseaux après apprentissage. En effet, l'apprentissage n'est arrêté que lorsqu'un minimum est atteint : l'arrêt est décidé en fonction de la valeur de la fonction de coût et de celle de la norme de son gradient.

Si le nombre de neurones est augmenté de façon incrémentale comme nous l'indiquons ci-dessus, on obtient forcément le prédicteur le plus parcimonieux réalisant la meilleure performance sans surajustement.

Il est à signaler que le prix à payer pour ces méthodes d'optimisation des réseaux de neurones est le temps de calcul. En général, on se contente dans la pratique d'utiliser des méthodes qui ne garantissent pas l'optimalité de la solution, mais ont l'avantage d'être plus rapides.

4.7.4 Le surajustement

Après cette revue des outils les plus fréquemment utilisés dans le domaine des réseaux de neurones pour sélectionner le meilleur modèle tout en évitant les solutions surajustées, les questions suivantes se posent :

- quelle est l'origine exacte du surajustement ?
- quelle démarche faudrait-il mettre en œuvre pour régler ce problème à la base ?

Le surajustement caractérise une fonction dont la complexité (c'est-à-dire le nombre et la nature des degrés de libertés) est telle qu'elle est capable de s'ajuster exactement aux exemples d'apprentissage, même si ceux-ci sont entachés de bruit.

Ce phénomène est donc, à l'origine, un phénomène local : dans certains domaines des entrées, la fonction utilise localement certains de ses degrés de liberté de manière à passer précisément par certains exemples.

Cette définition du surajustement suppose que tous les exemples ont la même importance et que l'on recherche effectivement une solution dont la réponse est satisfaisante.

Généralement, cette hypothèse est formalisée dans la fonction de coût choisie. Ainsi, pour éviter le surajustement, il faudrait limiter l'influence de chaque exemple sur l'estimation des poids du modèle. Les valeurs de ceux-ci doivent être déterminées par l'ensemble de la base d'apprentissage, et non par un exemple particulier.

4.7.5 Problème des minima locaux

Les minima trouvés par les algorithmes d'apprentissages sont souvent des minima locaux. Le minimum trouvé dépend du point de départ de la recherche, en d'autres termes de l'initialisation des poids. En pratique, il faut effectuer plusieurs minimisations avec des initialisations différentes, pour trouver plusieurs minima et retenir le « meilleur ».

Il est néanmoins impossible et généralement inutile, de s'assurer que le minimum choisi est le minimum global [IV.25].

4.8 Etude du pouvoir de généralisation des réseaux de neurones

La généralisation concerne la tâche accomplie par le réseau une fois son apprentissage achevé [IV.44]. Elle peut être évaluée en testant le réseau sur des données qui n'ont pas servi à l'apprentissage. Elle est influencée essentiellement par quatre facteurs : la complexité du

problème, l'algorithme d'apprentissage, la complexité de l'échantillon (le nombre d'exemples et la manière dont ils représentent le problème) et enfin la complexité du réseau (nombre de poids).

Parmi ces quatre facteurs, la complexité du réseau constitue un facteur important surtout dans les applications industrielles. En effet, lorsque l'on doit traiter des données bruitées, ce que nous supposons tout au long de cette thèse, l'objectif est de trouver le modèle optimal, présentant le meilleur compromis possible entre performance d'apprentissage et capacité de généralisation.

Le problème de la généralisation est souvent vu sous deux perspectives différentes. Dans la première, la taille du réseau est fixée (en accord avec la complexité du problème) et la question est : combien d'exemples d'apprentissage sont nécessaires pour atteindre une bonne généralisation ? Cette perspective est intéressante dans les applications où l'on a la possibilité d'acquérir autant d'exemples que l'on veut. Dans le second cas, le nombre d'exemples d'apprentissage est fixé et la question est : quelle taille du réseau donne la meilleure généralisation pour ces données ? On est conduit à adopter ce point de vue lorsque l'on est limité dans la possibilité d'acquérir des données d'apprentissage : il importe alors de déterminer quelle est la taille permettant au réseau de décrire au mieux les données en notre possession.

4.9 Validation et Tests

Alors que les tests concernent la vérification des performances d'un réseau de neurones hors échantillon et sa capacité de généralisation, la validation est parfois utilisée lors de l'apprentissage. Une fois le réseau de neurones développé, des tests s'imposent afin de vérifier la qualité des prévisions du modèle neuronal.

Cette dernière étape doit permettre d'estimer la qualité du réseau obtenu en lui présentant des exemples qui ne font pas partie de l'ensemble d'apprentissage. Une validation rigoureuse du modèle développé se traduit par une proportion importante de prédictions exactes sur l'ensemble de la validation. Si les performances du réseau ne sont pas satisfaisantes, il faudra, soit modifier l'architecture du réseau, soit modifier la base d'apprentissage.

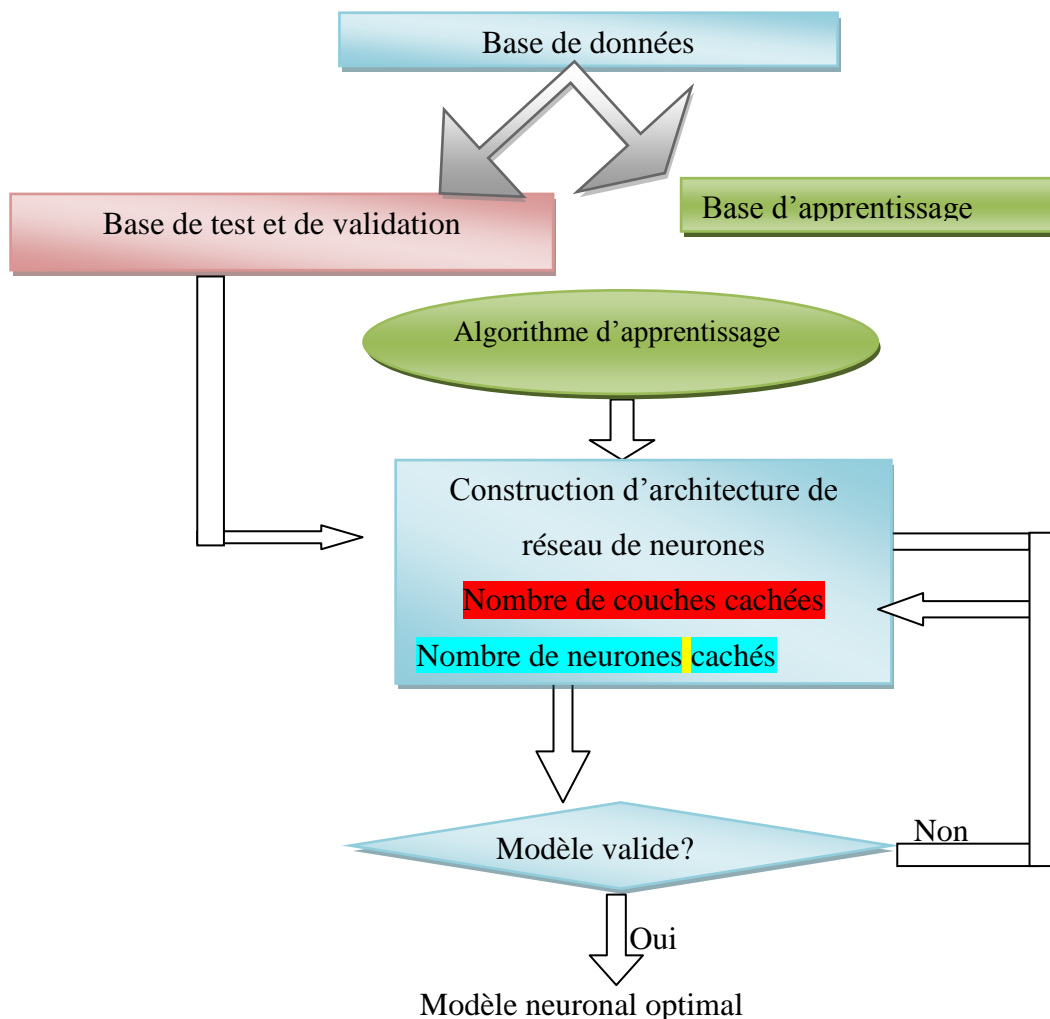


Figure 4.6: Organigramme de conception d'un réseau de neurones.

4.10 Conclusion

Après un aperçu historique sur les réseaux de neurones, et le rappel de quelques théorèmes relatifs à leurs capacités d'approximation, nous avons présenté un algorithme d'apprentissage très utilisé dans la communauté connexionniste : l'algorithme de rétro-propagation (RP). Plusieurs travaux lui ont été consacrés et un certain nombre de modifications ont été proposées pour accélérer le temps d'apprentissage et améliorer les performances du réseau obtenu. Nous en avons rappelé dans ce chapitre les principales variantes.

Dans la pratique, l'objectif d'une modélisation statistique n'est pas d'ajuster finement un modèle sur un ensemble d'apprentissage, mais d'obtenir un bon compromis entre performances d'apprentissage et performances de généralisation.

Lors de la construction d'un réseau de neurones, l'objectif donc est de déterminer la meilleure architecture possible et le meilleur jeu de poids au sens d'un critère donné (la généralisation par exemple).

Sous cet angle, la construction du réseau de neurones peut être vue comme un problème de recherche d'une solution dans l'espace des architectures et des poids.

Le prix à payer pour cette approche est le temps de calcul. En général, on se contente dans la pratique d'utiliser des méthodes qui ne garantissent pas l'optimalité de la solution, mais qui ont l'avantage d'être plus rapides.

Résumons à présent les points fondamentaux qu'il convient de toujours garder à l'esprit lorsque l'on cherche à mettre en œuvre des réseaux de neurones :

- les réseaux de neurones sont des outils *statistiques*, qui permettent d'ajuster des fonctions non linéaires très générales à des ensembles de points ; comme toute méthode statistique, l'utilisation de réseaux de neurones nécessite que l'on dispose de données suffisamment nombreuses et représentatives ;
- les réseaux de neurones artificiels à une couche cachée sont des approximateurs *parcimonieux* ;
- les réseaux de neurones permettent de modéliser des phénomènes statiques (réseaux non bouclés) et dynamiques (réseaux bouclés ou récurrents) ;
- il est toujours souhaitable, et souvent possible, d'utiliser, pour la conception des réseaux de neurones, les connaissances mathématiques dont on dispose sur le phénomène à modéliser : les réseaux de neurones ne sont pas nécessairement des « *boîtes noires* ».

Bibliographie

CHAPITRE IV

- [IV.1] **Fausset L.**, Fundamentals of Neural Networks. Architecture - Algorithms and Applications, 1st Edition, *Printice Hall, Inc.*, 1994.
- [IV.1] **Fausset L.**, Fundamentals of Neural Networks. Architecture - Algorithms and Applications, 1st Edition, *Printice Hall, Inc.*, 1994.
- [IV.2] **McCulloch W.** and **W. Pitts**, A logical calculus of the ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics*, Vol. 5, 1943, pp.115-133.
- [IV.4] **Rosenblatt F.**, The Perceptron: a Perceiving and Recognizing Automaton, *Project PARA, Report 85-460-1, Cornell Aeronautical Lab.*, 1957.
- [IV.5] **Rosenblatt F.**, The Perceptron: a probabilistic model for information storage and organization in the brain, *Psychological Review*, Vol. 65, 1958, pp.386-408.
- [IV.6] **Widrow B.** and **Hoff M.**, Adaptive switching circuits, *WESCON Convention Record, New York: IRE*, Vol. 4, 1960, pp.96-104.
- [IV.7] **Minsky M.**, et **Papert S.**, Perceptrons, *MIT Press, Cambridge, MA*, 1969.
- [IV.8] **Grosberg S.**, Adaptive Pattern Classification and Universal recoding, I : Parallel Development and coding of neural features detectors, *Biological Cybernetics*, Vol. 23, 1976, pp.121-134.
- [IV.9] **Wu J. K.**, Neural Networks and Simulation Methods, 1st Edition, *Marvel Dekker, Inc.*, 1994.
- [IV.10] **Hopfield J.J.**, Neural networks and physical systems with emergent collective computational abilities, *Proceedings of the National Academy of Sciences*, Vol. 79, 1982, pp.2554-2558.
- [IV.11] **Ackley D.**, **Hinton G.** and **Sejnowski T.**, A learning algorithm for Boltzmann machines *Cognitive Science*, Vol. 9, 1985, pp.147-169.
- [IV.12] **Rumelhart D.E.**, **Hinton G.E.**, and **Williams R.J.**, Learning Internal Representations by Error Propagation. In *Parallel Distributed Processing : Explorations in the Microstructures of Cognition*, Vol. 1, *D. Rumelhart and J. McClelland Eds. Cambridge: MIT Press*, 1986, pp.318-362.
- [IV.13] **Cybenko G.**, Approximation by superposition of a sigmoid function. *Math. of Control, Signals and Systems*, Vol. 2(4), 1989, pp.303-314.
- [IV.14] **Funahashi K.** , On the Approximate Realization of Continuous Mappings by Neural Networks, *Neural Networks*, Vol. 2, 1989, pp.183-192.

- [IV.15] **Hornik K.**, Stinchcombe M. and White H., Multilayer Feedforward Networks are Universal Approximators, *Neural Networks*, Vol. 2, 1989, pp.359-366.
- [IV.16] **Hornik K.**, Stinchcombe M. and White H. and Auer P., Degree of approximation results for feedforward networks approximating unknown mappings and their derivatives, *Neural Computation*, Vol. 6, 1994, pp.1262-1275.
- [IV.17] **Rivals I., Personnaz L.**, Dreyfus G., Ploix J.L., Modélisation, classification et commande par réseaux de neurones : principes fondamentaux, méthodologie de conception et applications industrielles, *Les réseaux de neurones pour la modélisation et la commande de processus*, J.P. Corriou éd., Lavoisier Tec et Doc, 1995.
- [IV.18] **Billings S.A.**, H.B., Jamaluddin H.B. and Chen S., Properties of Neural Networks With Applications to Modelling non-linear Dynamical Systems, *Int. Journal of Control*, Vol. 55, No 1, 1992, pp.193-224.
- [IV.19] **Dreyfus G.**, Idan Y., The Canonical Form of Non-linear Discrete-Time Models, *Neural Computation* Vol. 10, No. 1, 1998, pp.133-164.
- [IV.20] **Hornik K.**, Stinchcombe M. and White H., Universal approximation of an unknown mapping and its derivatives using multiplayer feedforward networks, *Neural Networks*, Vol. 3, 1990, pp.551-560.
- [IV.21] **Park J.** and Sandberg I.W., Universal Approximation Using Radial-Basis-Function Networks, *Neural Computation* Vol. 3, No. 2, 1991, pp.246-257.
- [IV.22] **White H.**, Artificial neural networks, Blackwell, New York, (1992).
- [IV.23] **Kohonen T.**, Self organized formation of topologically correct feature maps, *Biol Cybernetics*, Vol. 43, pp. 59-69, (1982).
- [IV.24] **Minsky M.**, et Papert S., *Perceptrons*, MIT Press, Cambridge, MA, 1969.
- [IV.25] **Bishop C.**, Neural Networks for Pattern Recognitions, *Clarendon Press-Oxford New-York*, 1995.
- [IV.26] **Werbos P.**, Beyond regression: new tools for prediction and analysis in the behavioural sciences, *PhD thesis, Harvard University, Cambridge, MA.*, 1974.
- [IV.27] **Park J.** and Sandberg I.W., Universal Approximation Using Radial-Basis-Function Networks, *Neural Computation* Vol. 3, No. 2, 1991, pp.246-257.
- [IV.28] **Le Cun Y.**, Une procédure d'apprentissage pour réseau à seuil asymérique. In *Cognitiva 85: A la Frontière de l'Intelligence Artificielle des Sciences de la connaissance des Neurosciences*, Paris: CESTA, Paris, 1985, pp.599-604.
- [IV.29] **Hassibi B.** and Stork D. G., Second order derivatives for network pruning : Optimal brain surgeon. In *S.J. Hanson, J.D. Cavnan, and C.L. Giles, Editors, Advances in Neural Information Processing Systems 5*, Morgan Kaufmann, San Mateo, CA, 1993, pp. 164-171.
- [IV.30] **Robbins H.** and Monro S., A stochastic approximation method, *Annals of Math. Stat.*, Vol.

22, 1951, pp.400-407.

- [IV.31] **Moussaoui A.K.**, Bouazza S.E. and Abbassi H.A., Feedforward Control of Hot Strip thickness Process Using Neural Networks, *Proceedings du Séminaire National sur l'Automatique et les Signaux, SNAS'99, U.B.M. Annaba, Algérie*, 1999, pp.237-241.
- [IV.32] **Gayner R.J.** and Downs T., On the properties of error functions that affect the speed of backpropagation learning, In *Marinaro M. and Morasso P. G., editors, ICANN'94, Proceeding of the International Conference on Artificial Neural Networks ; 26-29 May 1994, Sorrento, Italy*, 1994, pp.557-560.
- [IV.33] **Plaut D.C.**, Nowlan S.J. and Hinton G.E., Experiments on learning by back-propagation, *Technical Report CMU-CS-86-126, Carnegie Mellon University - Pittsburgh, PA 15213*, 1990.
- [IV.34] **Jacobs R.A.**, Increased rates of convergence through learning rate adaptive, *Neural Networks, Vol. 1*, 1988, pp.295-307.
- [IV.35] **Silva F.M.**, Almeida L.B., Speeding up backpropagation. In R. Eckmiller, editor, *Advanced Neural Computers, Elsevier Science Publishers B.V., North-Holland*, 1990, pp. 151-158.
- [IV.36] **Tollenaere T.**, SuperSAB: fast adaptive back propagation with good scaling properties, *Neural Networks, Vol. 3*, 1990, pp.561-573
- [IV.37] **Fletcher R.**, Practical methods of optimisation, *John Wiley and Sons Ltd, second edition, New-York, USA*, 1987.
- [IV.38] **Hertz J., Krogh A. and Palmer R.G.**, Introduction to the theory of neural computation, *Computation and neural systems series. Addison-Wesley, New-York, NY*, 1991.
- [IV.39] **Le Cun Y.**, Boser B., Denker J.S., Hendersen D., Howard R.E., Hubbard W., Jackel L.D., Backpropagation applied to handwritten zip code recognition. *Neural Computation, Vol. 1*, 1989, pp.541-551.
- [IV.40] **Bruenelli R.**, Training nets through stochastic minimization. *Neural Networks, Vol. 7(9)*, 1994, pp.1405-1412
- [IV.41] **Hagan M.T. and Mehaj M.** "Training feed-forward networks with the Marquardt Algorithm". *IEEE Transactions and Neural Networks*, 5(6), 1994, pp.989-993
- [IV.42] **Battiti R.** « First and second order methods for learning: Between steepest descent methods and newton's methods ». *Neural Computation, Vol. 4, No. 2*, 1992, pp. 141-166.
- [IV.43] **Sjöberg J.**, Hjalmeron H. and Ljung L., Neural Networks in System Identification, *Preprints 10th IFAC symposium on SYSID, Copenhagen, Denmark. Vol.2*, 1994, pp.49-71.

Chapitre 5

Application de réseaux de neurones pour la prédiction des mobilités dans les mélanges SF₆-N₂

5.1 Objectif

Ces dernières années ont vu les tentatives faites par un certain nombre d'auteurs d'utiliser divers (ANN) des modèles à base de réseaux de neurones artificiels pour aider à mieux prédire les tensions seuils des décharges couronnes dans les mélanges de SF₆[V.1-V.2- V.3] , Mais

Les applications des réseaux de neurones artificiels pour la prédiction des Mobilités des porteurs de charges dans les mélanges SF₆-N₂ sont très rares. Notre objectif est de construire un réseau de neurones artificiels capable de prédire des mobilités ioniques des porteurs de charges dans les mélanges SF₆-N₂. Ce modèle doit prédire la valeur de la mobilité ionique pour n'importe quelle pression et concentration de N₂ différente de celles utilisées pendant l'apprentissage. La prédiction du modèle est ensuite comparée avec l'expérimentation.

Nous scindons notre base de données en deux sous-ensembles : le premier est utilisé pour l'apprentissage, le second pour la validation de la réponse du réseau de neurones.

5.2 Construction du réseau de neurones

Lors de l'élaboration des RNA, plusieurs paramètres entrent en jeu comme les fonctions d'activation, les algorithmes d'apprentissage, le nombre de couches dans le réseau, etc.

La structure générale du modèle de prédiction neuronal proposé prend la forme représentée sur la figure 5.1.

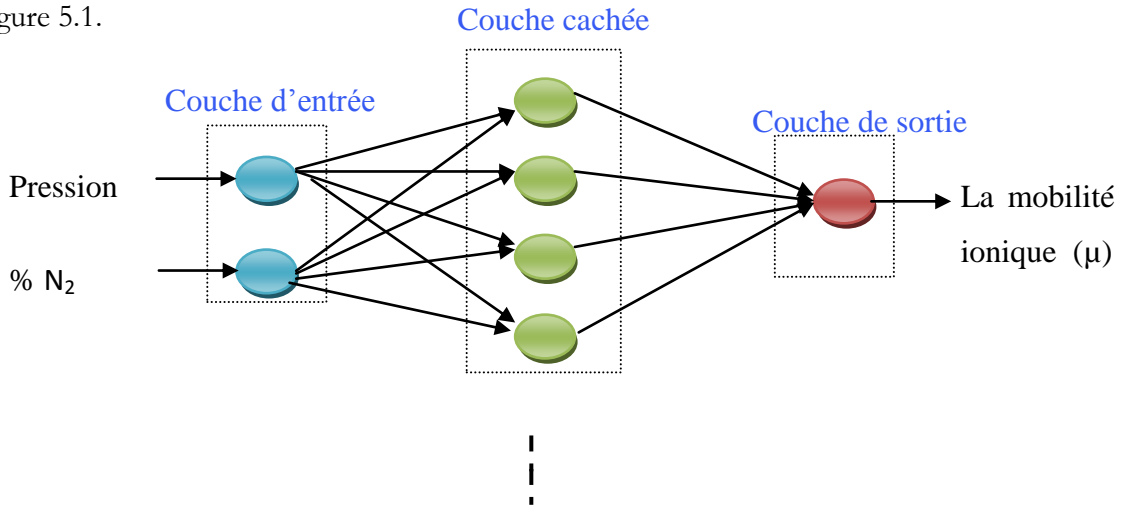


Figure 5.1:Schéma de la structure neuronale utilisée pour représenter la Mobilité des porteurs de charges

Le modèle neuronal est décrit par l'équation suivante :

$$\mu = \mathbf{NN}(\mathbf{P}, \mathbf{r}, \mathbf{w}) \quad (5.2)$$

Où : μ La mobilité ionique,

P: pression,

r: pourcentage du N_2 dans le mélange et w: le poids

Le modèle neuronal a comme entrées : la pression du gaz et le pourcentage du N_2 dans le mélange et comme sortie : les mobilités.

- La couche cachée renferme quatre neurones. Elle est utilisée pour améliorer la capacité d'apprentissage du réseau de neurones. Selon le théorème de Kolmogorov Hecht [V.4-V.5], le nombre de neurones dans la couche cachée doit être le double de neurones dans la couche d'entrée plus un neurone.
- L'apprentissage et les tests de ce réseau de neurones artificiel sont élaborés en utilisant le langage Matlab.

Afin d'effectuer un apprentissage précis, la couche de sortie est normalisée pour avoir le même ordre de grandeur de la couche d'entrée.

5.2.1 Choix de l'architecture

Tableau 5.1. Paramètres choisis

Paramètres	choix
Modèle utilisé	Perceptron multicouches (MLP)
Entrée	Pression, densité du gaz
Algorithme d'apprentissage	Levenberg-Marquardt
Nombre de couches	3
Fonction d'activation au niveau de la couche cachée	Sigmoïde (logsig)
Durée d'apprentissage	5000 epochs
Critère de performance	10^{-10}

Le modèle neuronal proposé est un perceptron monocouche, type feedforward utilisé pour la prédiction, avec un apprentissage supervisé (Tableau 5.1.) et un algorithme d'apprentissage selon la méthode de Levenberg-Marquardt.

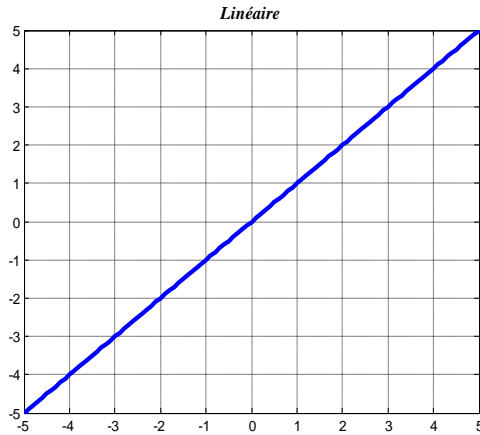
Cet algorithme se déroule comme suit :

- 1- Initialiser aléatoirement les poids du réseau
- 2- Présenter le premier exemple et calculer la sortie
- 3- Comparer la sortie avec la sortie réelle et calculer l'erreur
- 4- Rétropropager l'erreur et adapter les poids :
- 5- Si l'erreur nouvelle $>$ l'erreur ancienne $\Rightarrow W$ nouveau = W ancien + ΔW ,
règle d'apprentissage
- 6- Si l'erreur nouvelle $<$ l'erreur ancienne \Rightarrow Pas de modification de poids
- 7- Test si la performance désirée est atteinte \Rightarrow Arrêt
- 8- Sinon arrêt avec le nombre d'itérations (Cycles, epochs.)

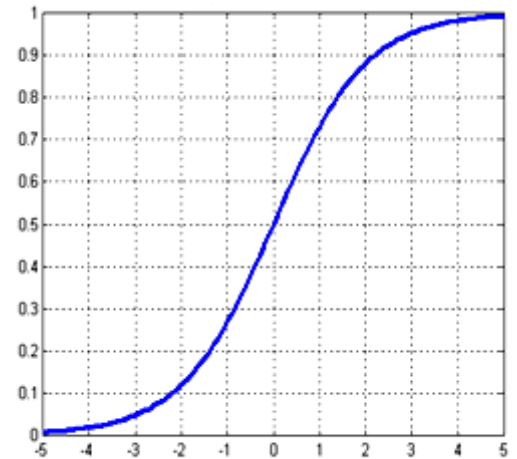
Où : W : vecteur des poids synaptiques

5.2.2 Choix de la fonction d'activation

La fonction d'activation des neurones de la couche cachée est une sigmoïde (cf. Chapitre 4), celle de la sortie est une fonction linéaire (figure 5.2); Ce qui donne au réseau de neurones proposé la propriété d'une approximation universel [V.6]. La combinaison linéaire des sigmoïdes permet d'approximer n'importe quelle fonction avec une précision arbitraire, à condition de disposer de suffisamment de neurones dans la couche cachée [V.7].



Sortie=purelin (Entée)
Fonction d'activation linéaire



Sortie=Logsig (Entée)
Fonction d'activation linéaire log-sigmoïde

Figure 5.2: Type de fonctions d'activation utilisées.

5.2.3 Algorithme d'apprentissage

Selon la figure 5.3, l'algorithme d'apprentissage ajuste les poids de toutes les liaisons de raccordement et les mobilités ioniques dans les nœuds de sorte que la sortie réelle $Y(t)$ et la sortie cible $T(t)$ sont réduites au minimum pour tous les motifs d'apprentissage donnés [V. 8]

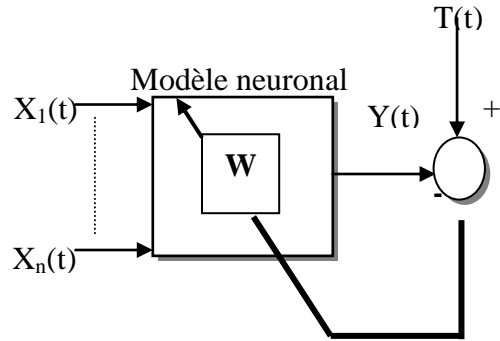


Figure 5.3: Schéma bloc de l'algorithme de prédiction

Pour les entrées d'apprentissage X ($i = 1 \dots n$), l'algorithme d'apprentissage est effectué pour minimiser la fonction d'énergie

$$E_p = \frac{1}{2} \sum_i [T_i - Y_i(N)]^2 \quad (5.3)$$

Où: $Y_i(N)$ est le $i^{\text{ème}}$ neurone de l'activation de la couche de sortie N , T_i est la $i^{\text{ème}}$ sortie désiré, et N est le nombre de couches. L'application de Levenberg -Marquardt à l'apprentissage de réseau neuronal est décrite dans [V.8-V.9]. Cet algorithme semble être la méthode la plus rapide pour l'apprentissage de réseaux de neurones multicouches. Les poids sont ajustés en fonction de la mise à jour itérative: Mise à jour des poids w_{ji} et w_{kj} selon une règle delta.

Pour permettre un ajustement en ligne des poids du modèle neuronal à chaque période d'échantillonnage moyennant une adaptation des poids selon un mode d'apprentissage incrémental.

Dans ce mode, les poids du modèle de prédiction sont ajustés selon l'équation

De mise à jour suivante :

Mise à jour des poids w_{ji} et w_{kj} selon une **règle Delta**

$$w^{(New)} = w^{(old)} + \Delta w$$

Avec

$$\Delta w = -\eta \frac{\partial J}{\partial w}$$

$$\mathbf{W}_{ij}^{(\text{new})}(\mathbf{n}) = \mathbf{W}_{ij}^{(\text{old})}(\mathbf{n}) - [\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}]^{-1} \mathbf{J}^T \mathbf{E}_T \quad (5.4)$$

Où: W_{ij} : Poids $i^{\text{ème}}$ neurone entre la couche $n + 1$ et $j^{\text{ème}}$ neurone de la couche n , \mathbf{I} est la matrice identité, \mathbf{J} est la matrice jacobienne, qui contient les dérivées premières des erreurs de réseau de neurones par rapport au poids et des biais et \mathbf{E}_T est un vecteur de toutes les erreurs de réseau ($0 \leq n \leq N-1$), Le paramètre μ est adapté au cours de la procédure d'apprentissage. Lorsque l'erreur totale E_T descend en dessous d'une valeur E_{\min} prédéterminée de la mobilité ionique, l'algorithme d'apprentissage est alors arrêté. La Retro-propagation est utilisée pour calculer le jacobien \mathbf{J} de la fonction de la performance E_p

Avec

$$\mathbf{J} = \frac{\partial E_p}{\partial \mathbf{W}_{ij}(\mathbf{n})} = \delta_i(\mathbf{n}) f'[u_i(\mathbf{n})] \cdot Y_j(\mathbf{n}-1) \quad (5.5)$$

Où:

$$\delta_i(\mathbf{n}) = \sum_j \delta_j(\mathbf{n}+1) f'[u_j(\mathbf{n}+1)] \cdot W_{ji}(\mathbf{n}+1) \quad (5.6)$$

$$\delta_i(\mathbf{N}) = -[T_i - Y_i(\mathbf{N})] \quad (5.7)$$

Avec $u_i(\mathbf{n})$: entrée au neurone $i^{\text{ème}}$ couche, L_{n+1} : nombre de neurones dans le $(n+1)^{\text{ème}}$ couche. $f'(\cdot)$ est la dérivée de la fonction d'activation sigmoïde sur l'entrée du neurone $j^{\text{ème}}$ à $(n+1)^{\text{ème}}$ couche. Il est important, cependant, de se rappeler que les modèles d'entraînement doivent couvrir toute la gamme de la combinaison d'entrée pour laquelle le réseau sera nécessaire pour effectuer l'émulation précise du processus.

La performance du réseau de neurones est évaluée en calculant l'erreur quadratique moyenne des vraies valeurs cibles des prédictions du réseau de neurones.

$$\text{MSE} = \frac{1}{N} \sum_{n=1}^N (t_n - y_n)^2 \quad (5.8)$$

Où y_n : sont les sorties du réseau

t_n : les valeurs cibles

N : le nombre d'échantillons.

5.3 Discussions des résultats

5.3.1 Apprentissage

En utilisant la base des données représentée sur la figure 3.7, la prédiction par les réseaux de neurones des valeurs des mobilités ioniques pour différentes pressions et différents taux du N_2 est plus simple et précise. Pour améliorer les performances de cette prédiction la base de l'apprentissage doit couvrir un large domaine, d'où la nécessité d'avoir une base de données plus étendue.

Sur la figure 5.4, Après la convergence du réseau de neurones proposé vers l'erreur quadratique moyenne désirée, ce qui démontre la réussite de la phase d'apprentissage du modèle ANN. Nous avons utilisé une nouvelle séquence de données, ayant les mêmes caractéristiques que la séquence d'apprentissage, pour effectuer un test de validation.

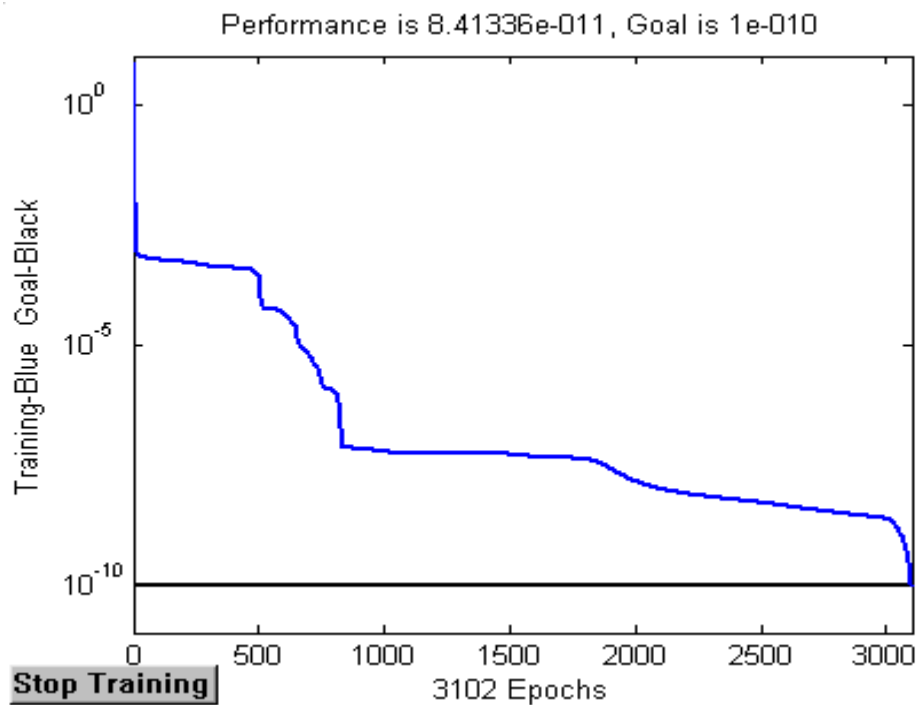


Figure 5.4: Evolution de l'erreur

Selon la courbe d'évolution du taux de l'erreur de prédiction donnée dans la Figure 5.5, la variance est sensiblement égale à sa valeur maximale en valeur absolue est inférieure à (10 %).

Normalisation des données Entrées/Sortie

La première Etant donné que les variables du RNA entrée et de sortie ont des ordres de grandeurs différents, l'alimentation de données d'origine sur le réseau de neurones, conduit à un problème de convergence. Il est évident que la sortie du RNA doit se situer dans l'intervalle de (0-1) et ce après la normalisation.

En outre, les signaux d'entrée doivent être conservés petites pour éviter un effet de saturation de la fonction sigmoïde. Ainsi, les données d'entrée-sortie sont normalisées avant l'apprentissage du réseau. La normalisation est effectuée en divisant les variables d'entrée-sortie à la valeur maximale des composants d'entrée et de sortie vectorielles. Les valeurs normalisées sont données par l'expression suivante:

$$\mu_N = a.\mu + b \quad (5.6)$$

Avec :

$$a = (\beta - \alpha) / (\max(\mu) - \min(\mu))$$

$$b = \beta - a.\max(\mu)$$

Les coefficients α et β sont choisis égaux à 0,1 et 1 respectivement.

La figure 5.5, présente le résultat pour le mélange SF₆-N₂. L'accord entre les mesures expérimentales et celles de la simulation est très marquant. L'allure des deux courbes est la même, il apparaît un écart réduit entre les deux courbes ce qui représente un résultat encourageant.

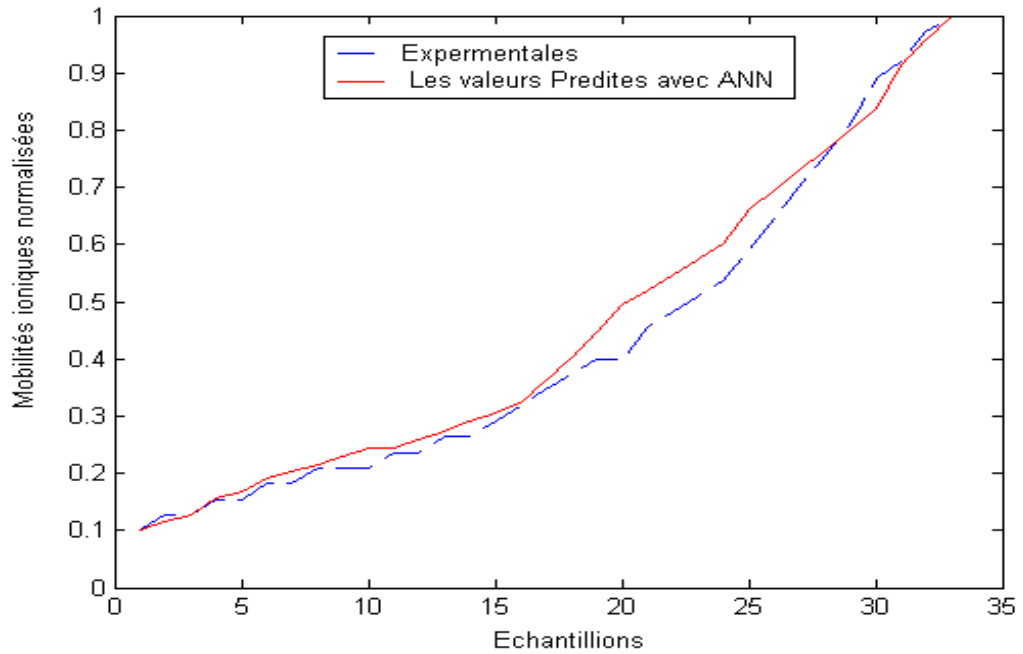


Figure 5.5: Mobilités ioniques mesurées normalisées comparés aux résultats obtenus par la prédiction des réseaux neurones à 4 bars en polarité négative

Selon la courbe d'évolution du taux de l'erreur de prédiction donnée dans la Figure 5.6, la variance est sensiblement égale à sa valeur maximale en valeur absolue est inférieure à (10 %).

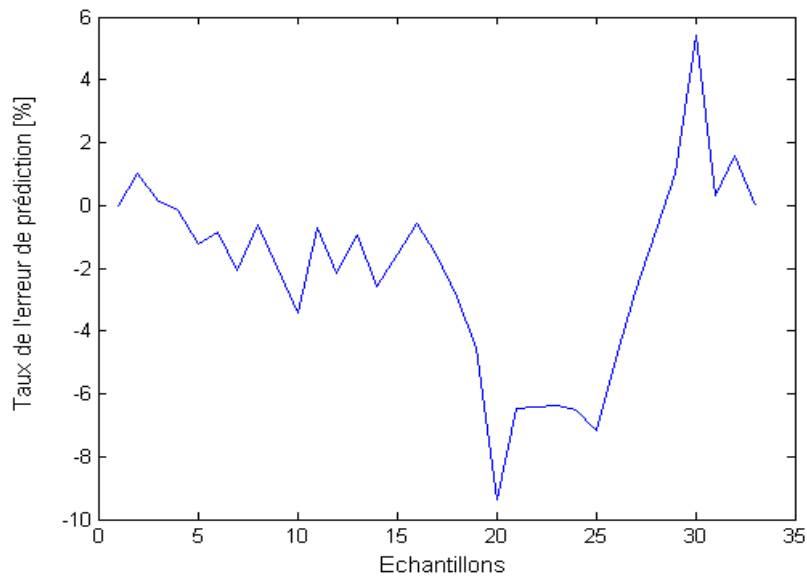


Figure 5.6: Evolution du taux de l'erreur de prédiction [%]

La figure 5.7, présente la comparaison entre la courbe validée des mobilités ioniques prédites par les réseaux neurones pour le mélange SF₆-N₂ et celle expérimentale à 4 bars en polarité négative. On voit qu'il y a une bonne concordance entre les valeurs expérimentales et celles validées.

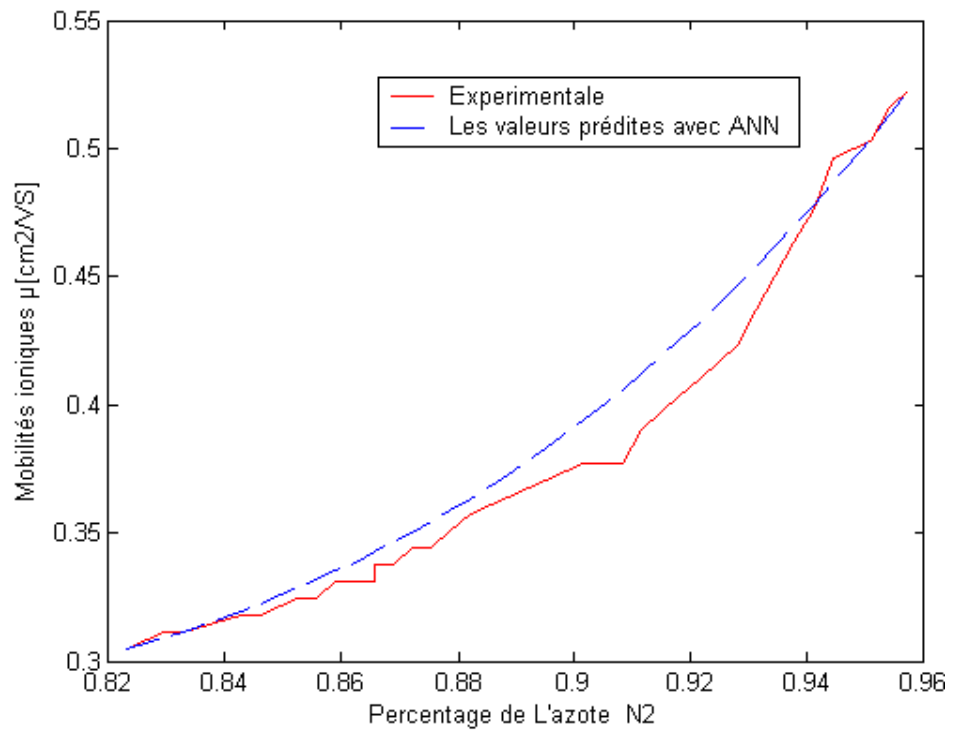


Figure 5.7: Mobilité ionique mesurée comparé au résultat obtenu par la prédiction des réseaux neurones à 4 bars en polarité négative

Sur la figure 5.8 on voit que le modèle peut prédire la valeur de la mobilité ionique pour n'importe quelle pression et concentration de N_2 dans le mélange SF_6-N_2 . Ceci est très important pour les professionnels du secteur de l'industrie de génie électrique.

Dans les mélanges gazeux SF_6-N_2 , les mobilités augmentent avec modération en fonction du pourcentage du N_2 (SF_6) dans le gaz jusqu'à une certaine limite du taux de Azote N_2 au delà duquel on assiste à une très forte augmentation des mobilités. On pense que ce phénomène est directement lié au changement de la nature des porteurs de charges, les ions O_2^- qui sont nettement plus légers, prennent le dessus sur les ions SF_6^- .

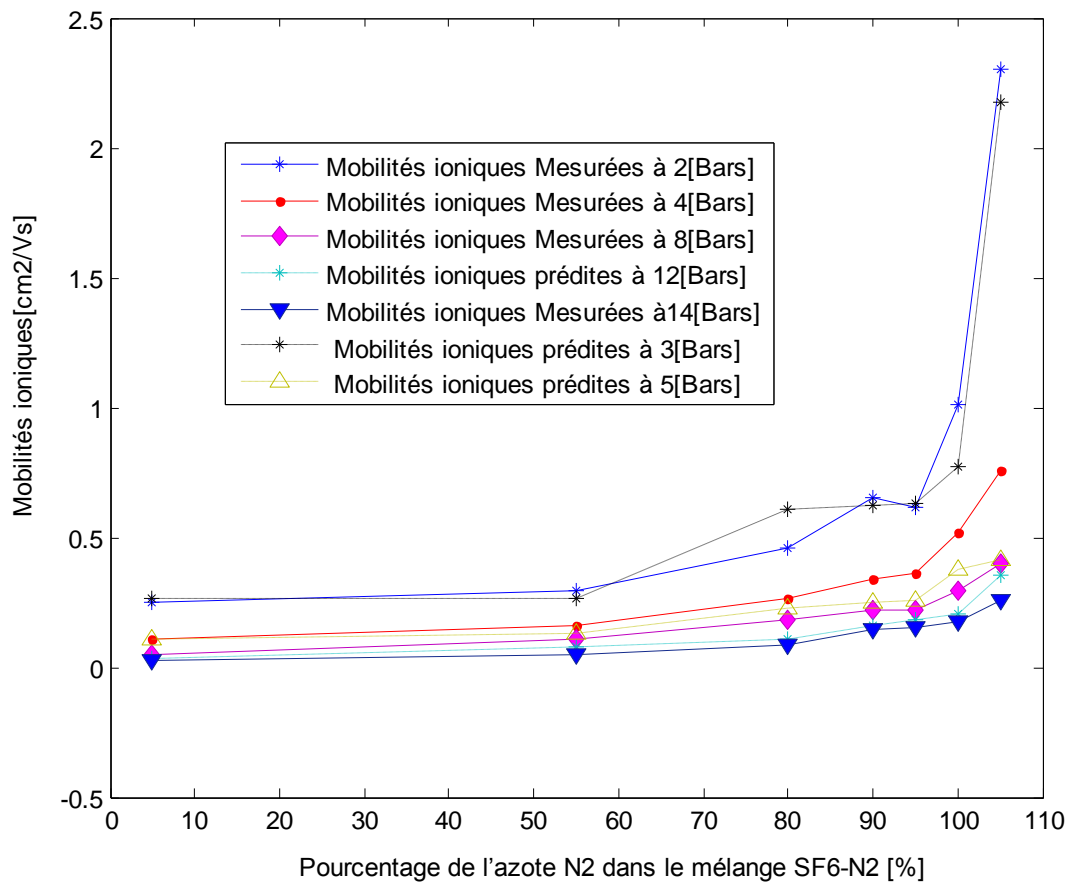


Figure 5.8: Mobilités prédites en fonction du pourcentage de l'azote N_2 dans le mélange SF_6-N_2 .

5.3.2 Résultats de simulation des mobilités ioniques

Des mesures de mobilité ionique ainsi que les simulations associées ont été réalisés sur quatre mélanges de gaz : SF₆-N₂ à 5% SF₆, à 10% SF₆, à 25% et à 100% SF₆. Pour chaque mélange sont présentées les courbes en polarité négative:

- Des mesures expérimentales sont comparées aux mesures prédites normalisées par les réseaux neurones courbes.
- De l'évolution évolution du taux de l'erreur de prédiction [%].
- Des mesures expérimentales sont comparées aux mesures prédites après validation.

Mélange SF₆-N₂ à 5% SF₆

La figure 5.9, présente la courbe normalisée des mobilités ioniques prédites par les réseaux neurones pour le mélange SF₆-N₂ à 5% de SF₆ et celle expérimentale. On voit qu'il y a une bonne convergence entre les valeurs expérimentales et celle normalisée. L'évolution du taux de l'erreur de prédiction reste dans les limites très acceptables comme on peut le voir sur la figure 5.10.

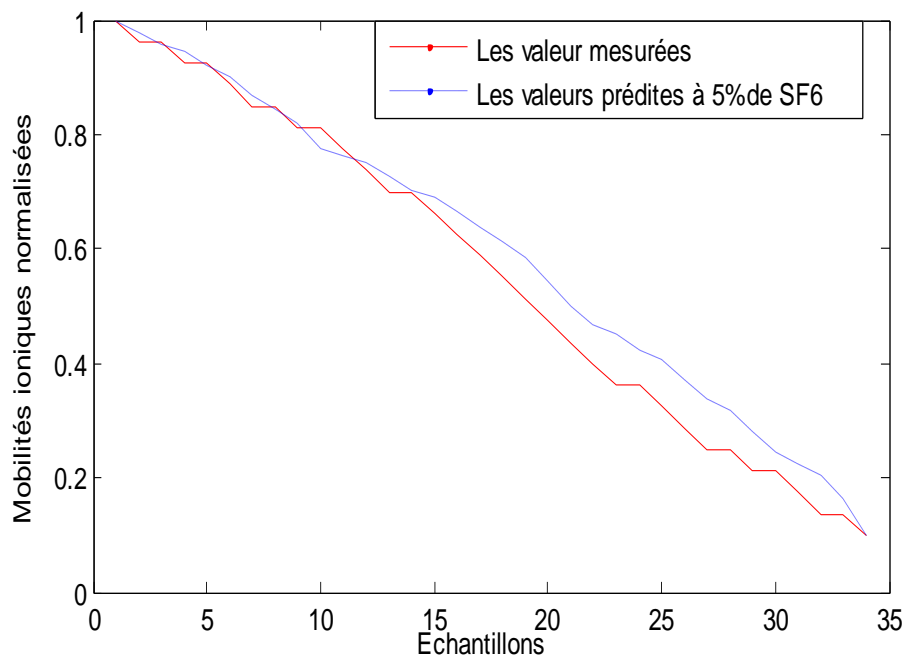


Figure 5.9:présentent la courbe normalisée des mobilités ioniques prédites par les réseaux neurones pour le mélange SF₆-N₂ à 5% de SF₆

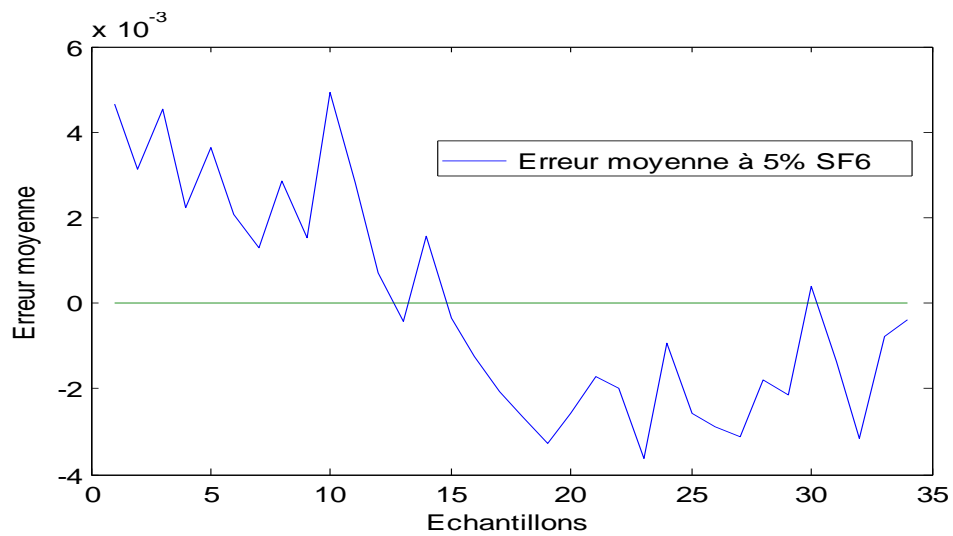


Figure 5.10: Evolution du taux de l'erreur de prédiction à 5% SF₆

La figure 5.11, présente la comparaison entre la courbe validée des mobilités ioniques prédites par les réseaux neurones pour le mélange SF₆-N₂ à 5% de SF₆ et celle expérimentale. On voit qu'il y a une bonne concordance entre les valeurs expérimentales et celles validées.

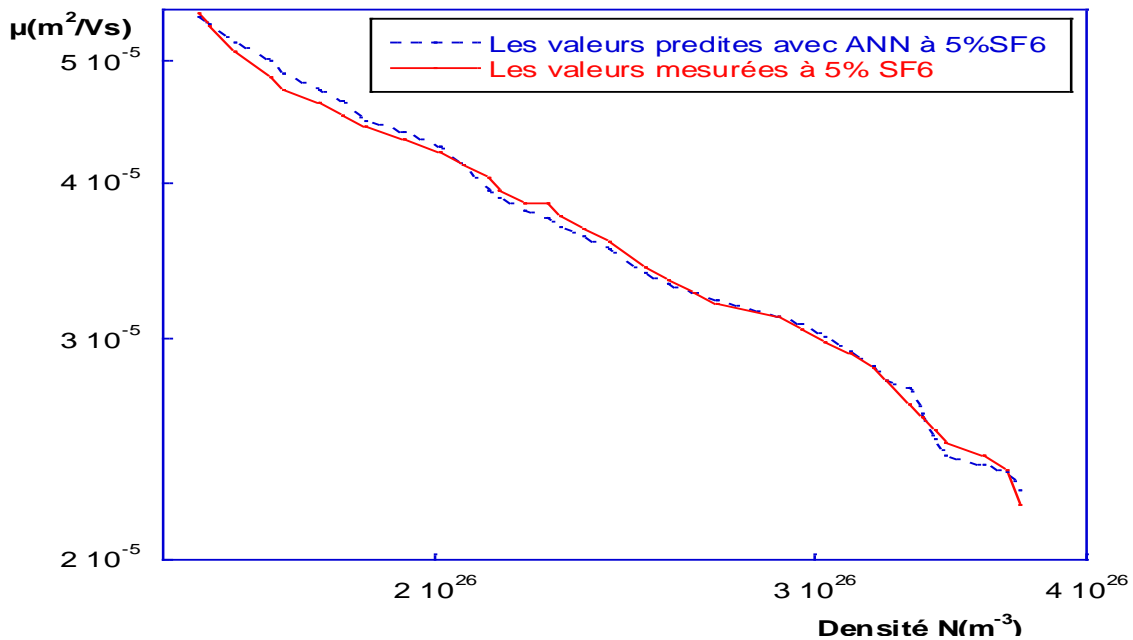


Figure 5.11: Mobilités ioniques mesurées et prédites pour un mélange à 5% de SF₆ en polarité négative.

Mélange SF₆-N₂ à 10% SF₆

Sur la figure 5.12 on peut observer que les mobilités ioniques prédites pour le mélange SF₆-N₂ à 10% de SF₆ en utilisant la prédiction par l'ANN pour les valeurs normalisées semblent être en bonne accord avec ceux obtenus expérimentalement.

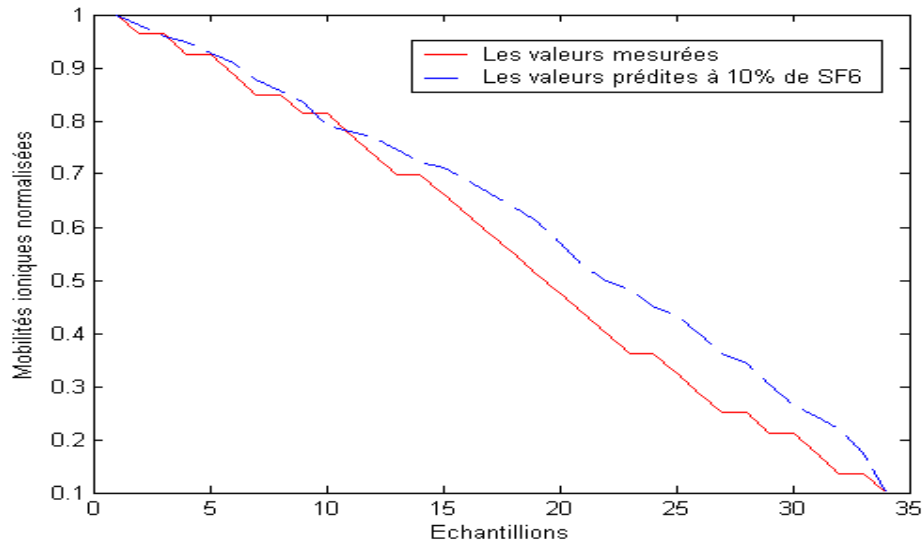


Figure 5.12: Mobilités ioniques mesurées et prédites (normalisées) pour un mélange à 10% de SF₆ en polarité négative.

L'évolution du taux de l'erreur de prédiction est présentée sur la figure 5.13 et sa valeur est inférieure à 10%.

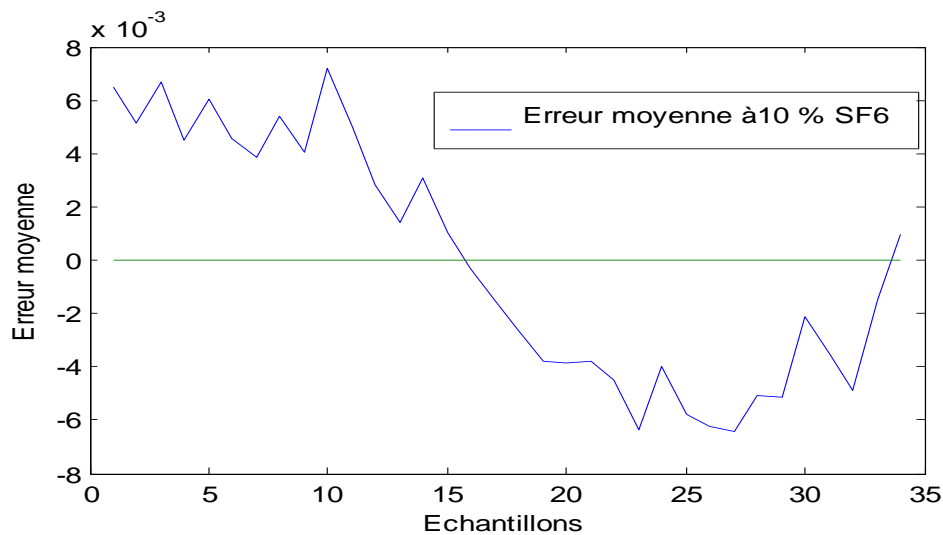


Figure 5.13: Evolution du taux de l'erreur de prédiction à 10 % SF₆

Après validation des mesures de prédiction par l'ANN la comparaison entre la courbe validée des mobilités ioniques et celle expérimentale pour le mélange SF₆-N₂ à 5% de SF₆ représentée sur la figure 5.14, donne des résultats très proche l'une par rapport à l'autre.

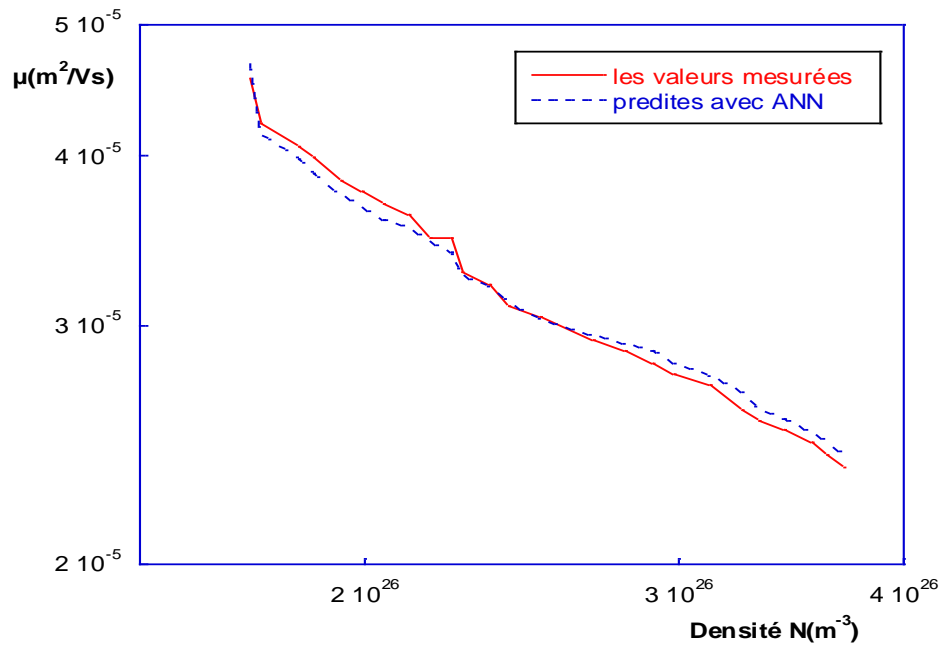


Figure 5.14: Mobilités ioniques mesurées et prédites pour un mélange à 10% de SF₆ en polarité négative.

Mélange SF₆-N₂ à 25% SF₆

La courbe des valeurs prédites normalisées pour le mélange SF₆-N₂ à 25% de SF₆ sont peu décalées par rapport à celle des valeurs mesurées comme on peut le voir sur la figure 5.15.

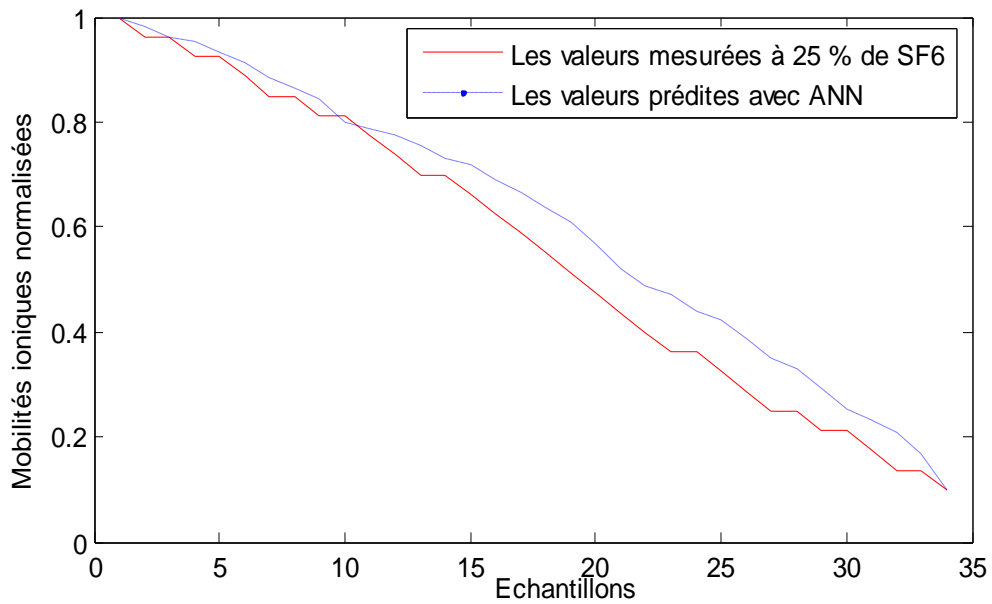


Figure 5.15: Courbes normalisées des mobilités ioniques prédites par les réseaux neurones pour le mélange SF_6-N_2 à 25% de SF_6 en polarité négative.

L'erreur présentée sur la figure 5.16 ne dépasse pas les 10% pour le mélange SF_6-N_2 à 5% de SF_6 .

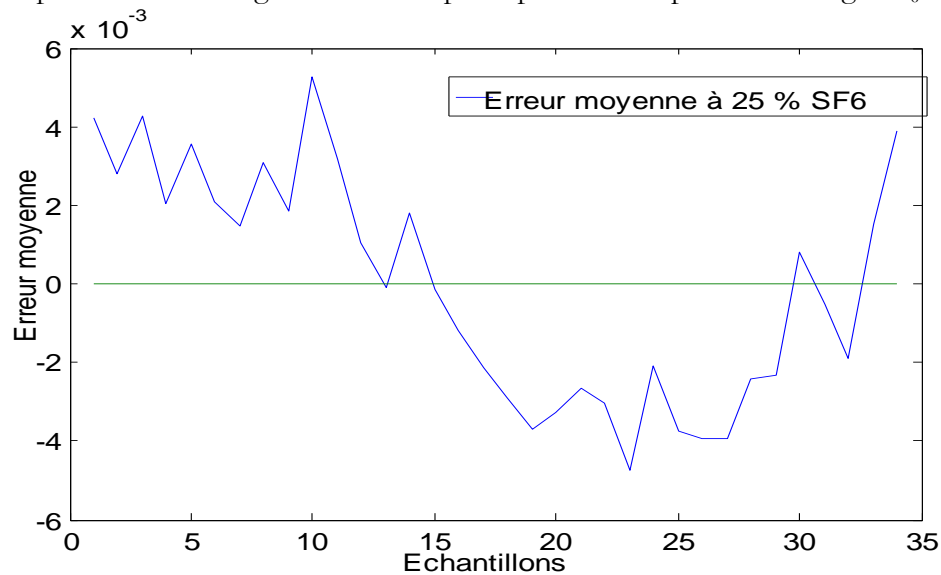


Figure 5.16: Evolution du taux de l'erreur de prédiction à 25% SF_6

Les valeurs des mobilités ioniques prédites et validées comparées aux mesures expérimentales pour le mélange SF_6-N_2 à 25% de SF_6 représentée sur la figure 5.17 sont légèrement différentes pour les faibles pressions par contre, pour les pressions élevées on constate que les résultats sont presque confondus.

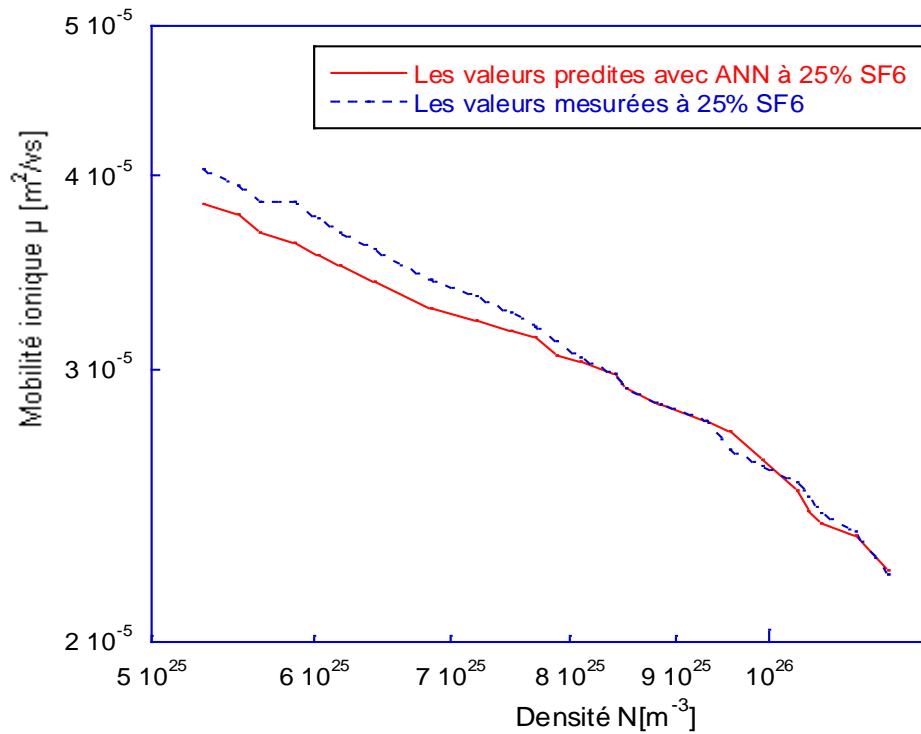


Figure 5.17: Mobilités ioniques mesurées et prédites pour un mélange à 25 % de SF₆ en polarité négative.

Mélange SF₆-N₂ à 100% SF₆

Les valeurs prédites normalisées et mesurées expérimentalement pour le mélange SF₆-N₂ à 100% de SF₆ sont tracées sur la figure 5.18. Les deux courbes présentent une concordance presque parfaite.

Le taux d'erreur es relativement petit come on peut le voir sur la figure 5.19.

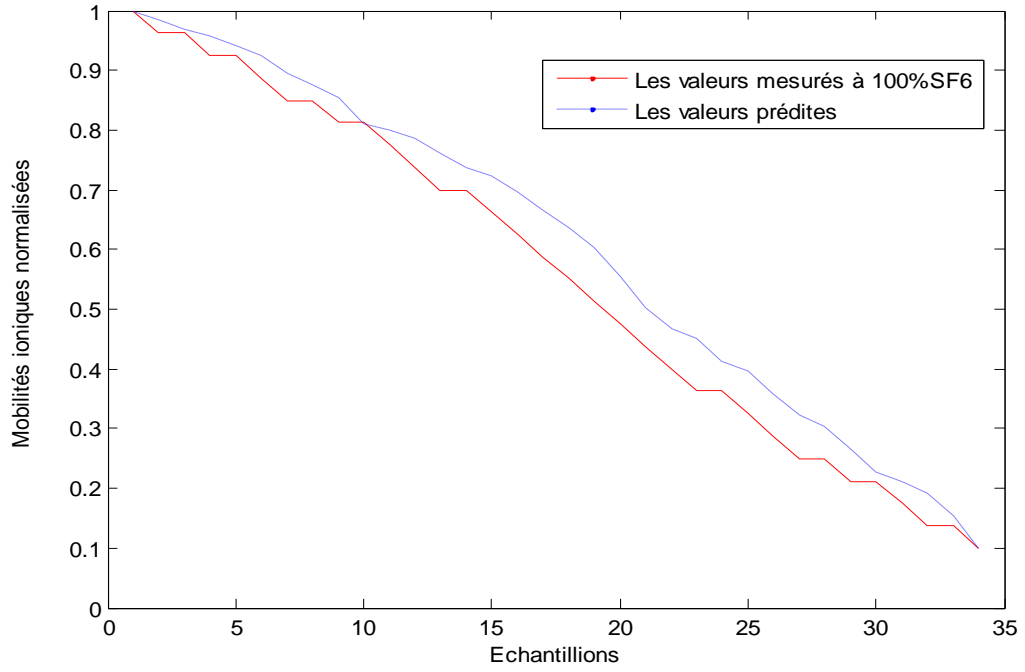


Figure 5.18: Mobilités ioniques normalisées prédites par les réseaux neurones pour le mélange SF_6-N_2 à 100% de SF_6

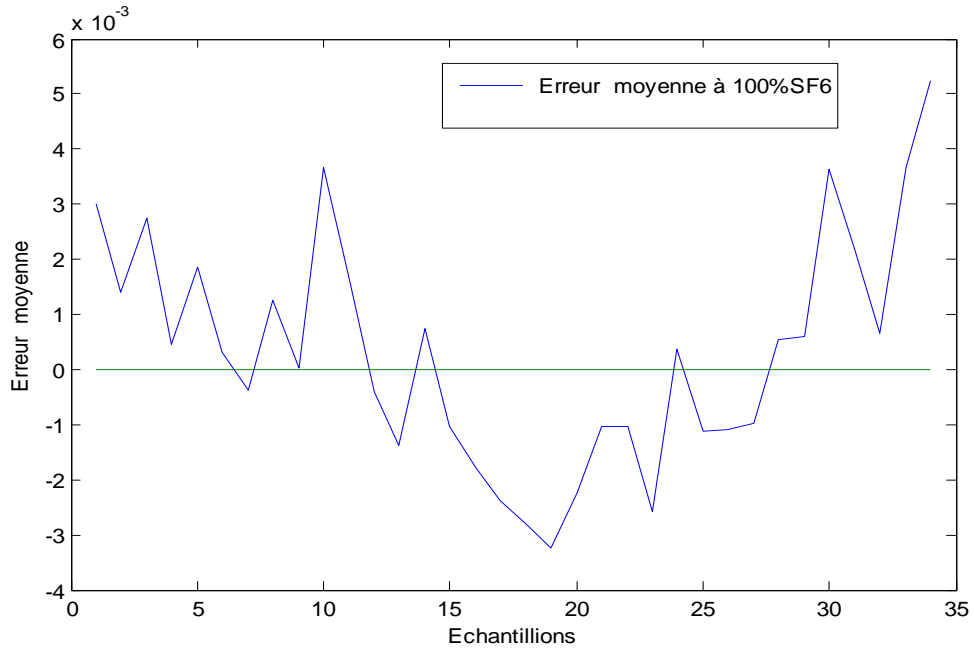


Figure 5.19: Evolution du taux de l'erreur de prédiction à 100% SF_6

La figure 5.20, présente la comparaison entre la courbe validée des mobilités ioniques prédites par les réseaux neurones pour le mélange SF₆-N₂ à 100% de SF₆ et la courbe expérimentale. On voit qu'il y a un bon accord entre les deux courbes.

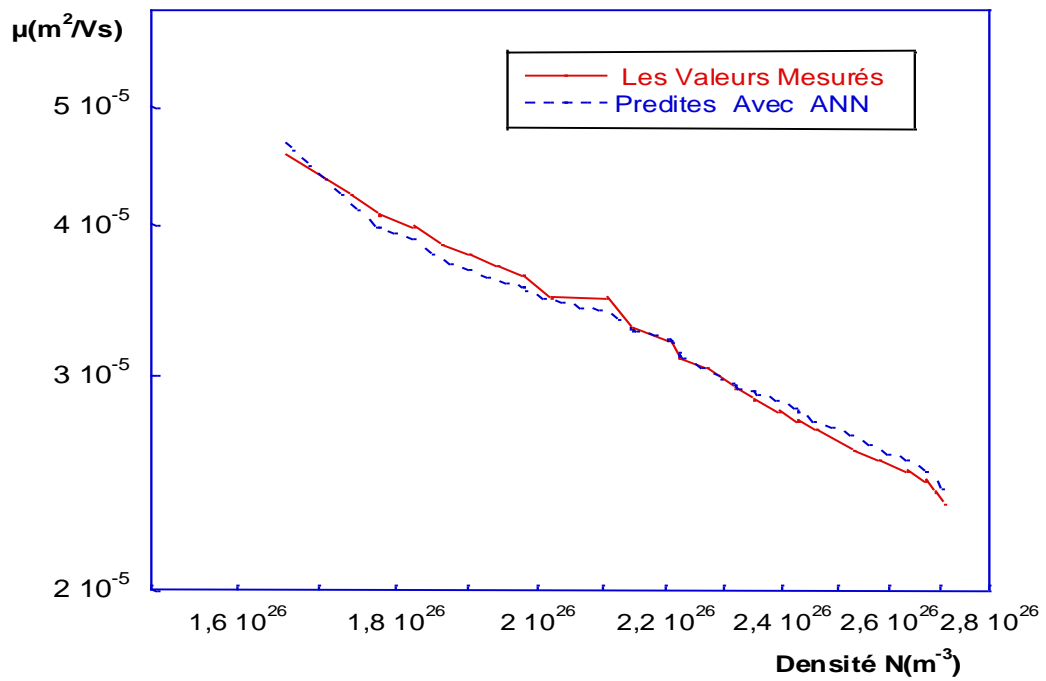


Figure 5.20: Mobilités ioniques mesurées et prédites pour un mélange à 100% de SF₆ en polarité négative.

Sur la figure 5.21 sont représentées les courbes des mobilités des porteurs de charges déterminées par la méthode indirecte (model de Sigmond), calculées par le théorie de Langevin, mesurées par W. F. Schmidt et finalement prédites par l'ANN pour le SF₆pur en polarité négative.

On constate qu'il y a des écarts entre les trois modèles, les modèles empiriques prédisent des valeurs de mobilités ioniques très proches au modèle de réseaux de neurones. L'écart entre les mobilités de Langevin et celles mesurées peut être attribué la prédominance à des pressions élevées aux ions plus lourds comme les clusters.

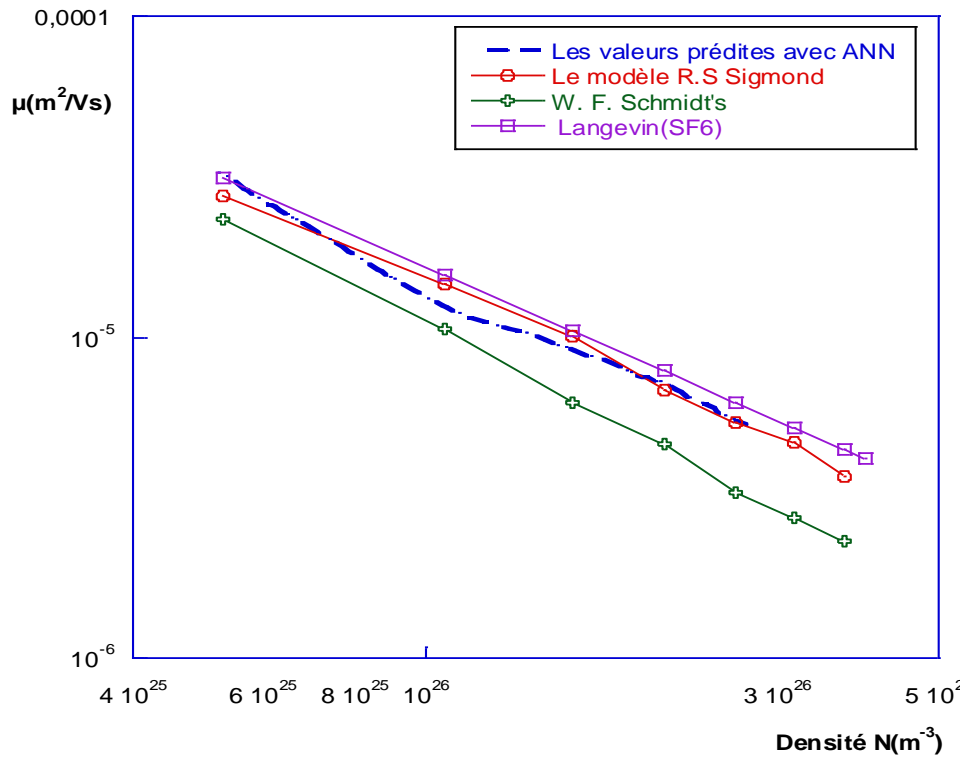


Figure 5.21: Courbes des mobilités des porteurs de charges déterminées par les différents modèles et celles déterminées Langevin comparées et celles mesurées par W. F. Schmidt aux résultats obtenus par la prédiction des réseaux neurones pour le SF₆ en polarité négative

La figure 5.22 représente la Variation de la mobilité en fonction de la densité pour différentes pressions. Nos résultats suivent toujours une variation en N^{-1} .

Nous obtenons un bon accord entre les valeurs expérimentales et les valeurs prédites, on peut constater que les mobilités du SF_6 nettement inférieures à celles du mélange à 5% de SF_6 et 90% de N_2 pour la même densité ceci explique pour les clusters $(SF_6)_2 SF_6^-$ qui sont beaucoup plus lourds que les ions N^- et N^+ qui se trouvent dans le mélange.

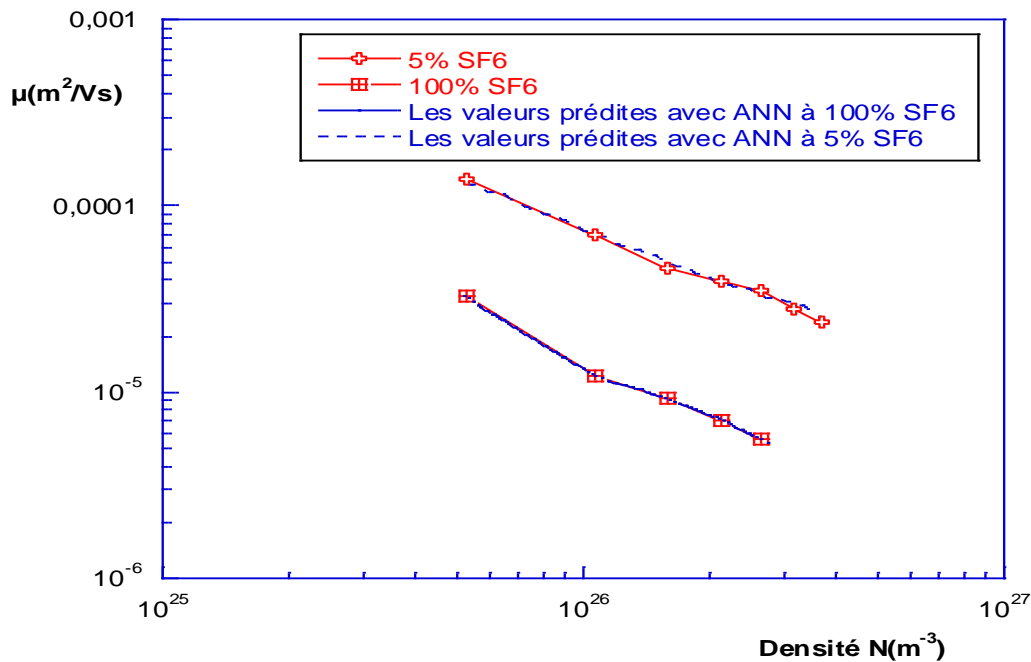


Figure 5.22: Variation de la mobilité en fonction de la densité pour différentes pressions

5.4 Conclusion

Les résultats obtenus des mobilités ioniques en utilisant le modèle empirique sont en concordance avec ceux déterminés par le modèle neuronal pour les mélanges SF_6-N_2 .

Parmi les conditions à satisfaire pour réaliser un apprentissage raisonnable du modèle neuronal de prédiction, on peut citer :

- ✓ la sortie du modèle doit être proche de celle du processus réel.
- ✓ l'erreur de prédiction finale doit être la plus faible possible.

L'outil de modélisation utilisé, en l'occurrence les réseaux de neurones, semble être un outil très efficace pour la prédiction des mobilités ionique dans le SF₆ et les mélanges SF₆-N₂.

L'avantage réel de l'utilisation des RNA pour la prédiction est la facilité d'implémentation. Une fois le bon choix effectué, il ne reste que la vérification de la reproductibilité du modèle établi par un test de généralisation.

Les RNA ne demandent aucun prétraitement ni simplification des valeurs d'entrée : Flexibilité et parcimonie vérifiées. Mais, sans une base de données bien homogène et dense les avantages des RNA peuvent être affectés.

Avec les résultats ainsi obtenus, nous espérons aboutir au but de ce travail :

- Construction d'un réseau de neurone à une couche cachée type *feed-forward* ayant les capacités de prédire des Mobilités des porteurs de charges pour les mélanges SF₆-N₂
- Apprendre à programmer avec le MATLAB : Dans MATLAB, l'élaboration d'un modèle neuronal semble être aisée du moment où l'application est claire (choix de l'échantillon d'apprentissage, type du RNA, nombre de neurones dans la couche cachée.
- Bien que le modèle empirique soit en accord avec les valeurs mesurés et ce pour les faibles pressions, il est remarqué leurs insuffisance pour les pressions élevés, ce qui justifie l'emploi des modèles de prédictions neuronaux qui sont performants pour le cas des pressions faibles et pression élevés

Bibliographie

- [V.1] S.S. Tezcan , M.S. Dincer , H.R. Hiziroglu, “Prediction of breakdown voltages in $N_2 + SF_6$ gas mixtures” Annual Report Conference on Electrical Insulation and Dielectric Phenomena 2006.
- [V.2] B. Monchusi, S. Letlotla H. Ilgner, J. McGill, “Modeling of Breakdown Voltage by Artificial Neural Network” 4th Robotics and Mechatronics Conference of South Africa (Robmech 2011), CSIR Pretoria South Africa, 23-25 November 2011.
- [V.3] Neelakantan, R. and J. Guiver, 1998. “Applying neural networks Hydrocarbon Processing”, pp: 91-96. Nabney, I.T., 2002.
- [V.4] Hornik, K., M. Stinchcombe and H. White, 1989. “Multilayer feed forward networks are universal approximations. Neural Networks”, 2: 359-366.
- [V.5] S. Bolat, “Electrode and Insulator Contour Optimization by Artificial Neural Networks”, MSc. Thesis, Institute of Science and Technology, Istanbul Technical University, 2003.
- [V.6] Cybenko G., Approximation by superposition of a sigmoid function. Math. of Control, Signals and Systems, Vol. 2(4), 1989, pp.303-314.
- [V.7] PARIZEAU M. Réseau de neurones ; (2004) Université LAVAL, pp.49-50.
- [V.8] S. Weerasooriya, M. A. Sharkawi, “Identification and Control of a DC Motor Using Back-Propagation Neural Networks”, IEEE Trans. On Energy Conversion, 6(4), 1991, pp.363-369. Discharges» IEEE Trans on dielect and Elect Insul .page(s) 618 - 622 vol.1, 2000.
- [V.9] S. Haykin, “Neural Networks: a Comprehensive Foundation”, Upper Saddle River, N.J., Prentice Hall, 1999.

Conclusion générale

L'Hexafluorure de soufre est depuis longtemps largement utilisé comme gaz d'isolation dans les appareillages moyenne et haute tension (disjoncteurs, postes sous enveloppe métallique, câble à isolation gazeuse...). Ce gaz possède d'excellentes qualités techniques telles qu'une bonne tenue diélectrique, Toutefois, il a été mis en évidence que le SF₆ est un gaz à effet de serre notable. Une solution pour le long terme consiste à son élimination complète ce qui est actuellement pratiquement impossible. La solution la plus plausible pour le court terme, est de réduire la concentration du SF₆ dans l'atmosphère en utilisant le mélange SF₆-N₂.

L'association de N₂ et de SF₆ permet d'utiliser au maximum les qualités de chacun des gaz. En effet, le SF₆ est un gaz électronégatif dont les capacités d'attachement sont particulièrement efficaces. L'azote, quant à lui, a de bonnes qualités de ralentissement des électrons énergiques. Ainsi dans le mélange les électrons aptes à ioniser les atomes et à déclencher une avalanche électronique, sont ralentis par les molécules d'azote et arrivent dans des domaines d'énergie où le SF₆ a les capacités d'attachement les plus grandes.

L'utilisation des caractéristiques courant-tension ($I = f(U)$) pour la détermination des mobilités des porteurs de charges dans le SF₆ et le mélange SF₆-N₂, montrent que les valeurs des mobilités sont inversement proportionnelles à la densité du gaz et augmentent avec la diminution du pourcentage du SF₆ dans le mélange SF₆-N₂. Cette méthode indirecte, utilise le modèle de Sigmond pour déterminer les paramètres de transport des porteurs de charges.

Pour le SF₆ pur, l'ion SF₆⁻ est prédominant mais à pressions élevées la formation des ions plus lourds comme les clusters est très probable, ce qui explique la différence observée entre la courbe théorique de Langevin et celle mesurée par Schmidt.

La prédiction des mobilités des porteurs de charges par la méthode de la technique des réseaux de neurones à donner des résultats très satisfaisant, en effet, les écarts entre les courbes obtenues par le méthode indirecte et celles obtenues en utilisant la simulation sont très minimales. Par conséquent, cette technique de simulation peut être préconisée pour la prédiction des paramètres de conduction dans les diélectriques gazeux.