



TRAITEMENT DU SIGNAL ET SES APPLICATIONS

Nice 7 au 12 mai 1973

DETECTION DANS UN BRUIT D'UN SIGNAL
PERTURBE PAR UN FILTRAGE LINEAIRE ALEATOIRE

Lucien SOURROUILLE

Laboratoire d'Etude des Phénomènes Aléatoires - Bât. 210
UNIVERSITE DE PARIS-SUD - Centre d'ORSAY 91405 ORSAY.**RESUME**

On établit la structure du "meilleur" récepteur pour deux types de situations :

- (a) on connaît la loi a priori du filtrage
- (b) on ignore cette loi et on fait appel uniquement aux propriétés de linéarité.

On trace les caractéristiques opérationnelles de réception de ces récepteurs et on les compare à celles du récepteur empirique (filtre adapté - quadrature - sommation) d'une part et du récepteur qui connaîtrait la forme du signal à chaque épreuve d'autre part.

SUMMARY

We derive the "best" detector in two kinds of situations :

- (a) we know the a priori probability of the filter
- (b) we are unaware of this probability and we only use the properties of linearity.

We plot the operating characteristics of this two detectors and of two others : an empiric detector (matched filter - squarer - sum) and a detector which would know the signal shape in each trial.

I - INTRODUCTION -

Nous traitons ici un problème de détection statistique qui apparaît notamment dans les théories des communications, du radar ou du sonar ; il s'agit de déceler, dans une observation qui comporte en permanence un bruit de fond, la présence éventuelle d'un signal additif. L'observation $x(t)$ est constituée par la donnée sur un intervalle de temps I d'une certaine grandeur physique (électromagnétique, acoustique, ...), enregistrée par exemple par une antenne de réception. Le bruit $B(t)$ est par définition une manifestation fluctuante, aléatoire, de cette grandeur ; l'observation $x(t)$ devra par conséquent être considérée comme une réalisation d'une fonction aléatoire $X(t)$.

Le signal est une manifestation de la grandeur physique en question, caractérisée sur I soit par sa donnée (signal certain), soit par ses propriétés statistiques (signal aléatoire) ; sa présence modifie la loi de $X(t)$.

Nous nous intéressons au cas important en pratique (voir par exemple [1] , [2]) où l'incertitude sur le signal éventuellement reçu $R(t)$ provient d'un filtrage linéaire aléatoire d'un signal émis $s(t)$ certain (nous appelons filtre aléatoire un filtre dont la réponse impulsionnelle $H(t)$ est une fonction aléatoire telle que la réponse à $s(t)$ existe presque sûrement)

II - FORMALISATION DU PROBLEME -

a) Nous ferons une approche temporelle et considérerons des grandeurs échantillonnées dans le temps avec un pas Δt . Nous noterons $f_i = f(i \Delta t)$, i entier, pour toute fonction f .

b) Nous adopterons comme intervalle d'observation $I = [0, T + D]$ où $T = n \Delta t$ est la durée du signal originel s , $D = p \Delta t$ est un majorant de la portée de la réponse impulsionnelle H du filtre aléatoire (ayant en vue des applications physiques, nous nous limiterons à des filtres causaux).



c) Le signal à détecter s'exprime alors par l'équation de convolution discrète :

$$(1) \quad R_i = \sum_{j=0}^p s_{i-j} H_j \quad i = 0, 1, \dots, n+p$$

où $H = \begin{bmatrix} H_0 \\ \vdots \\ H_p \end{bmatrix}$ est un vecteur aléatoire à valeurs dans \mathbb{R}^{p+1} .

Sous forme matricielle (1) s'exprime par :

$$(2) \quad R = M H \quad \text{où } M = \begin{bmatrix} s_0 & 0 & & & \\ s_1 & s_0 & & & \\ \vdots & \vdots & \ddots & & \\ s_n & s_{n-1} & \dots & s_0 & \\ 0 & & & & s_n \end{bmatrix} \text{ matrice } (n+p+1) \times (p+1)$$

d) Nous supposons le bruit B indépendant de H , centré, gaussien, blanc, de puissance a^2 , c'est-à-dire que :

$$(3) \quad E [B_i B_j] = a^2 \delta_{ij} \quad i, j = 0, 1, \dots, n+p.$$

Nous nous trouvons alors en présence d'hypothèses statistiques à tester : comment distinguer, à partir d'une réalisation du vecteur aléatoire X , entre

$$(\mathcal{H}_0) \quad X = B \text{ (bruit seul), et}$$

$$(\mathcal{H}_1) \quad X = B + MH \text{ (signal présent) ?}$$

La loi de probabilité de X est indexée par le paramètre θ appartenant à un ensemble Θ (ici, $\Theta: \mathbb{R}^{p+1}$). Dans un problème de test d'hypothèses, on s'intéresse à une partition de la famille $\{P_\theta\}$ des lois de X , c'est-à-dire à une partition de Θ en Θ_0 , correspondant à l'hypothèse \mathcal{H}_0 , et Θ_1 , correspondant à l'hypothèse \mathcal{H}_1 .

Prendre une décision d_0 (resp. d_1) concernant X , c'est choisir comme vraie l'hypothèse \mathcal{H}_0 (resp. \mathcal{H}_1).

Une fonction de décision (ou stratégie, ou test) est une application S de \mathcal{X} (espace d'observation = ensemble des valeurs possibles de x) dans $\Delta = \{d_0, d_1\}$.

La théorie de la décision, en introduisant une

perte $L(h, d)$ encourue lorsqu'on prend la décision d alors que h est la vraie valeur du paramètre, permet de définir l'optimalité d'une stratégie.

III - LE RECEPTEUR OPTIMAL POUR UNE LOI DE PROBABILITE DE H DONNEE :

Le paramètre H (dont les réalisations sont notées h) étant aléatoire et de loi connue, on peut définir pour toute stratégie S un risque moyen :

$$\mathcal{R}_S = E \left[L(H, S(X)) \right]$$

Le critère de Bayes consiste à choisir S de façon à minimiser \mathcal{R}_S .

Lorsqu'on dispose des lois de probabilité de H dans \mathcal{H}_0 et \mathcal{H}_1 , mais qu'on ignore les probabilités a priori des deux hypothèses, le critère de Bayes est inapplicable, mais on peut alors adopter le point de vue de Neyman-Pearson : parmi tous les tests de niveau inférieur ou égal à un nombre α donné ($0 < \alpha < 1$) on choisira celui de puissance maximale †.

Dans les problèmes de test d'hypothèses, où la perte $L(h, d)$ ne dépend que de l'appartenance de h à \mathcal{H}_0 ou à \mathcal{H}_1 , les points de vue de Bayes et de Neyman-Pearson conduisent au même résultat 3,4, à savoir :

Comparer le rapport de vraisemblance défini par $\Lambda(x) = \frac{P_1(x)}{P_0(x)}$ = rapport des densités de probabilité de X sous \mathcal{H}_1 et \mathcal{H}_0 à un seuil K ; suivant que $\Lambda(x)$

est supérieur (resp. inférieur) à K , décider $S(x) = d_1$ (resp. d_0), K étant déterminé soit par les probabilités a priori des hypothèses \mathcal{H}_0 et \mathcal{H}_1 (Bayes), soit par le niveau α du test (Neyman-Pearson).

Remarques : - l'hypothèse \mathcal{H}_0 est simple dans notre problème, puisque \mathcal{H}_0 se réduit à un seul élément $h = 0$;

† Dans les problèmes de détection le niveau d'un test = $P [S(X) = d_1 | \mathcal{H}_0]$

est communément appelé probabilité de fausse-alarme ; la puissance = $P [S(X) = d_1 | \mathcal{H}_1]$ est appelée probabilité de détection correcte.



- l'hypothèse \mathcal{H}_1 étant composite, le calcul du rapport de vraisemblance, c'est-à-dire de $P_1(x)$, nécessite la connaissance de la loi de H sous \mathcal{H}_1 ; nous pouvons prendre pour cette loi une distribution sur tout \mathcal{D} , à condition que $P[\mathcal{D}_0] = P[H = 0] = 0$.

a) Envisageons le cas d'une loi gaussienne pour H , de moyenne μ et de covariance Γ , ce que nous noterons $H = \mathcal{N}[\mu, \Gamma]$.

comme $B = \mathcal{N}[0, a^2 I]$ est indépendante de H : MH est aussi gaussienne = $\mathcal{N}[M\mu, M\Gamma M^t]$;

$$B + MH = \mathcal{N}[M\mu, M\Gamma M^t + a^2 I].$$

Ceci se traduit pour les densités de probabilité par :

$$(4) \quad P_0(x) = \exp - \frac{1}{2a^2} x^t x,$$

$$(5) \quad P_1(x) = \exp - \frac{1}{2} (x - M\mu)^t (M\Gamma M^t + a^2 I)^{-1} (x - M\mu).$$

Mettons $\Lambda(x)$ sous la forme :

$$(6) \quad \Lambda(x) = \exp \left\{ - \frac{1}{2a^2} (x - \eta)^t (I-A) (x - \eta) + C \right\}$$

où c ne dépend pas de x et où nous avons posé

$$(7) \quad A = \left(\frac{M\Gamma M^t}{a^2} + I \right)^{-1} ;$$

on obtient (6) avec

$$(8) \quad \eta = - \left(\frac{M\Gamma M^t}{a^2} \right)^{-1} M\mu.$$

Alors, comparer $\Lambda(x)$ à un seuil revient à comparer à un autre seuil la fonction monotone de $\Lambda(x)$ suivante :

$$(9) \quad (x - \eta)^t (I-A) (x - \eta) ;$$

Soit, en posant $y = M^t x$ et $G = M^t M$:

$$(10) \quad T_1 = (y + a^2 \Gamma^{-1} \mu)^t (G + a^2 \Gamma^{-1})^{-1} (y + a^2 \Gamma^{-1} \mu).$$

La structure quadratique tient au caractère gaussien de l'observation sous chacune des deux hypothèses [5, 6] .

L'introduction de la nouvelle variable aléatoire $Y = M^t X$ se justifie pleinement car, d'après la structure (2) de M , M^t représente la sortie du filtre adapté au signal s , prise en compte à partir de $t = T$, lorsque x en est l'entrée.

h étant un vecteur de \mathbb{R}^{p+1} , les valeurs $r = Mh$ du signal à détecter forment un sous-espace vectoriel E de dimension $p+1$ de l'espace d'observation \mathbb{R}^{n+p+1} . L'incertitude sur le signal tient à ces $p+1$ degrés de liberté ; le filtre adapté a pour rôle de réduire la dimension des données de $n+p+1 = m$ (variable X) à $p+1 = \dim E$ (variable Y).

En définitive, on décidera suivant :

$$T_1 \begin{cases} \geq & K \quad d1 \\ < & \quad do \end{cases}$$

où T_1 correspond à la structure :

$$\begin{aligned} & \text{FILTRE ADAPTE - FORME QUADRATIQUE } (G + a^2 \Gamma^{-1})^{-1} \\ & = \text{FILTRE ADAPTE - QUADRATION - SOMMATION PONDEREE.} \end{aligned}$$

b) Notre modèle de signal nous permet également de retrouver un résultat déjà établi [7] pour un signal à temps d'arrivée aléatoire :

il suffit de considérer pour H les valeurs $\begin{bmatrix} \delta_{ok} \\ \delta_{lk} \\ \vdots \\ \delta_{pk} \end{bmatrix}$ (δ = Symbole de Kronecker) correspondant aux différents retards $k\Delta t$, avec les probabilités q_k .

Le test basé sur le rapport de vraisemblance

$$\Lambda(x) = \frac{\int_{\Theta} P_h(x) q(h) dh}{P_0(x)}$$

aboutit alors à comparer à un seuil la statistique correspondant à la structure :

FILTRE ADAPTE - OPERATION NON LINEAIRE (exponentiation) - SOMMATION PONDEREE (par les q_k).

IV - LE RECEPTEUR INVARIANT -

Lorsqu'on ne connaît pas la loi de probabilité du paramètre H , on ne peut définir d'optimalité absolue pour une



procédure de décision.

Comme il n'existe en général pas de test uniformément plus puissant, on est amené à restreindre la classe des tests en concurrence et à chercher parmi ceux-là celui qui domine uniformément les autres.

Nous utiliserons uniquement la linéarité des hypothèses à tester, qui fait que notre problème de détection reste invariant par rapport à certaines transformations sur l'espace d'observation \mathcal{X} .

Rappelons [8] qu'un groupe G de transformations sur \mathcal{X} laisse invariant un problème de décision statistique si :

(i) G laisse invariante la famille $\{P_h\}$:
pour tout $g \in G$, pour tout $h \in \mathbb{H}$ il existe $h' \in \mathbb{H}$ tel que si P_h est la loi de X , alors $P_{h'}$ est la loi de $g(X)$.
Si h' est unique : G induit sur \mathbb{H} un groupe \bar{G} de transformations $\bar{g} : h' = \bar{g}(h)$.

(ii) G et \bar{G} laissent invariante la fonction de perte $L(h, d)$:
pour tout $g \in G$, pour tout $d \in \Delta$, il existe $d' \in \Delta$ tel que
 $L(h, d) = L(\bar{g}(h), d')$ pour tout $h \in \mathbb{H}$.

Cette invariance de L induit sur Δ un groupe \tilde{G} de transformations $\tilde{g} : d' = \tilde{g}(d)$.

Une stratégie S est invariante par rapport à G si :

$$(11) \quad S(g(x)) = \tilde{g}(S(x)) \text{ pour tout } x \in \mathcal{X}, \text{ pour tout } g \in G.$$

Dans un problème de test d'hypothèses où

$$(12) \quad L(h, d) = C_{ij} \text{ si } h \in \mathbb{H}_i, d = d_j ; i, j = 0, 1,$$

la condition d'invariance amène, pour le cas général où les C_{ij} sont distincts, à ne considérer que des transformations telles que

$$(13) \quad \bar{g}(\mathbb{H}_i) = \mathbb{H}_i \text{ et } \tilde{g}(d) = d.$$

Une stratégie invariante est alors telle que

$$(14) \quad S(g(x)) = S(x).$$

Dans notre problème, G est le groupe des transformations orthogonales qui conservent E (il faut en effet trouver pour tout

$x = Mh + k$, un h' tel que $y(x) = Mh' + g(b)$, avec $y(B)$ ayant la même loi que B ; d'où le résultat).

On adoptera donc, pour fonder la stratégie invariante (14) une statistique invariante par rapport à G_1 .

La statistique

(15) $T_2 = \| \text{projection de } X \text{ sur } E \|^2$

répond à la question.

T_2 ayant un rapport de vraisemblance monotone vis-à-vis du paramètre h , on montre que le test

(16) $T_2 \underset{<}{\underset{K}{\geq}} d_1$

est uniformément plus puissant parmi les tests invariants de son niveau $[9]$.

Compte-tenu de la définition de E , T_2 prend la forme :

(17) $T_2 = \gamma^t G^{-1} \gamma$

on retrouve ici aussi la structure

FILTRE ADAPTE - QUADRATION - SOMMATION

Notons que les récepteurs définis par T_1 et T_2 coïncident dans le cas d'un signal d'amplitude aléatoire (c'est-à-dire dans le cas où $\dim E = 1$).

Remarque : l'invariance fournit également une solution au problème de détection dans le cas où la puissance du bruit est inconnue. Le test est alors basé sur une statistique T invariante par rapport aux transformations orthogonales qui conservent E et par rapport aux homothéties. Soit sur la comparaison à un seuil de la quantité :

$$T = \frac{\| \text{Projection de } X \text{ sur } E \|^2}{\| \text{distance de } X \text{ à } E \|^2}$$

V - CARACTERISTIQUES OPERATIONNELLES DE RECEPTION

(C.O.R.) :

Considérons un détecteur basé sur la comparaison à



un seuil K d'une statistique $T(X)$. Les C.O.R. sont les courbes β versus α définies paramétriquement par :

$$\alpha_k = P [T > K | \mathcal{H}_0] = \text{probabilité de fausse alarme}$$

$$\beta_k = P [T > K | \mathcal{H}_1] = \text{probabilité de détection correcte.}$$

Outre les récepteurs T_1 et T_2 établis précédemment, nous introduisons :

- Le récepteur T_3 défini empiriquement par un filtrage adapté (au signal primitif) suivi de quadrature-sommation.

La statistique T_3 s'écrit

$$(18) \quad T_3 = Y^t Y.$$

- Le récepteur T_4 assimilé à un "diable" qui connaîtrait à chaque épreuve la forme du signal à détecter. Conditionnellement à une valeur du paramètre h de filtrage, T_4 réalise un filtre adapté ; il faut ensuite considérer la moyenne de ses performances suivant la loi adoptée pour H .

Prenons comme loi de H la loi gaussienne pour laquelle T_1 est optimal.

Nous avons établi les C.O.R. de ces 4 récepteurs moyennant quelques hypothèses supplémentaires permettant de pousser les calculs jusqu'au bout [10] :

- la v. a. H est supposée centrée, de matrice de covariance $\Gamma = \sigma^2 I$;

- on admet qu'on dispose, pour prendre la décision, de deux épreuves indépendantes.

Une première série de courbes (fig. 1 à 4) a été bercée par des dimensions de \mathcal{X} et Γ fixées, mais en faisant varier le rapport signal à bruit d'entrée (SBE).

L'écart de T_4 aux autres récepteurs permet d'évaluer la perte due à l'incertitude sur la forme du signal.

T_1 , optimal pour la loi gaussienne choisie, est évidemment au-dessus de T_2 et T_3 , mais plus ou moins loin suivant SBE.

Lorsque SBE croît, T_2 passe au-dessus de T_3 et se rapproche de T_1 ; ce fait peut-être interprété de la façon



suiivante : SBE grand signifie, à puissance du bruit constante, une grande dispersion de valeurs de H ; c'est-à-dire une loi a priori sur H peu informative.

Une deuxième série de courbes (fig. 5 à 7) a été tracée avec SBE constant, mais en faisant varier la dimension du sous-espace E engendré par le signal : on vérifie que T_1 , T_2 , T_3 sont très proches pour $\dim E$ faible, puis s'écartent lorsque $\dim E$ croît.

CONCLUSION

Il vaut mieux, évidemment, utiliser la loi a priori des coefficients H du filtre, ... si on la connaît et si on peut la mettre en oeuvre.

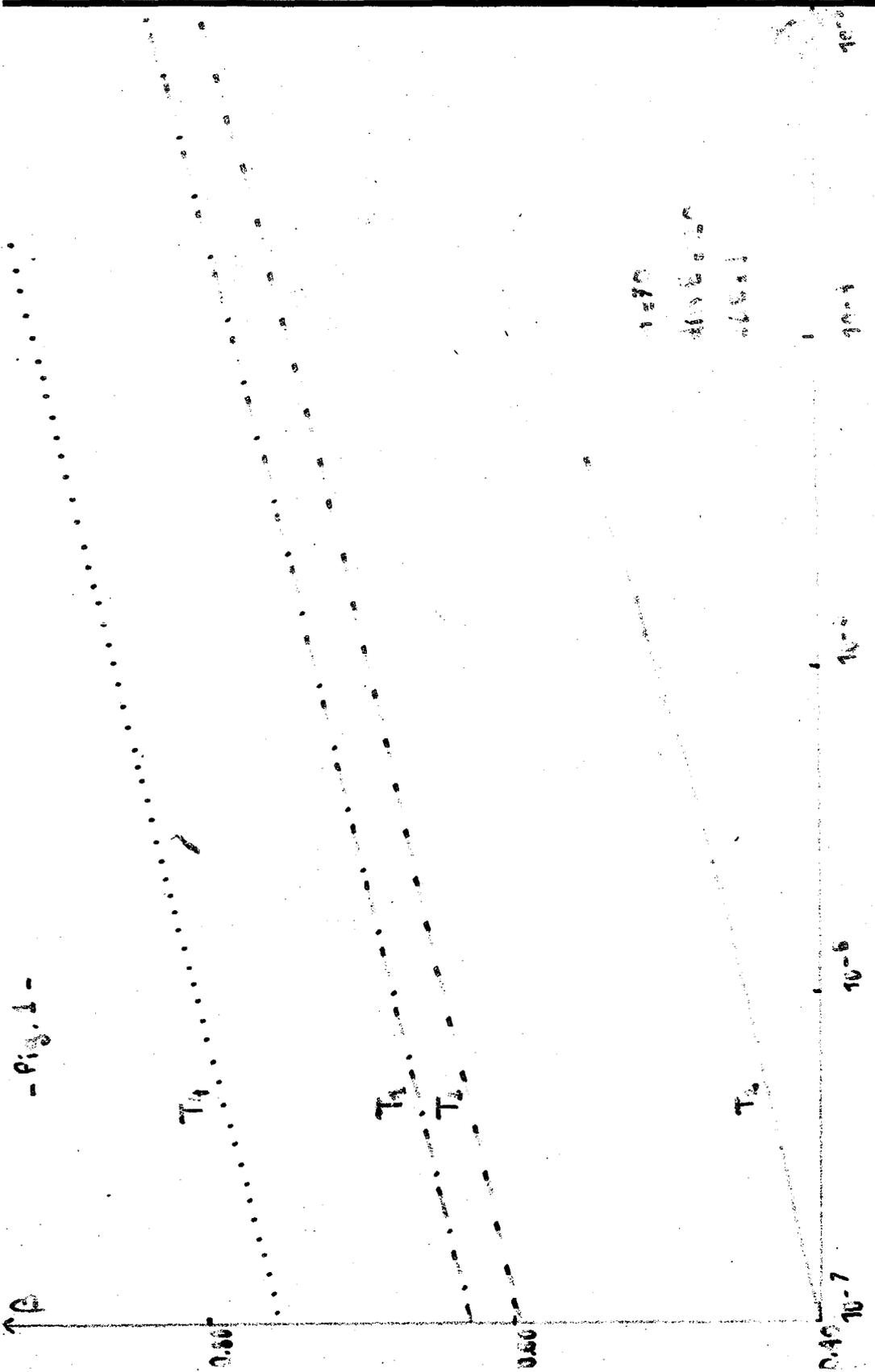
Le récepteur "empirique" T_3 (filtre adapté - quadrature-sommation), peut donner des résultats acceptables lorsqu'on ne connaît pas la loi a priori de H .

Le récepteur invariant T_2 ne semble fournir de bons résultats que pour un rapport signal sur bruit d'entrée élevé, ce qui est peu intéressant dans la plupart des situations pratiques; mais il demande encore à être testé contre T_3 , notamment quand la loi réelle du paramètre H s'écarte du modèle gaussien.

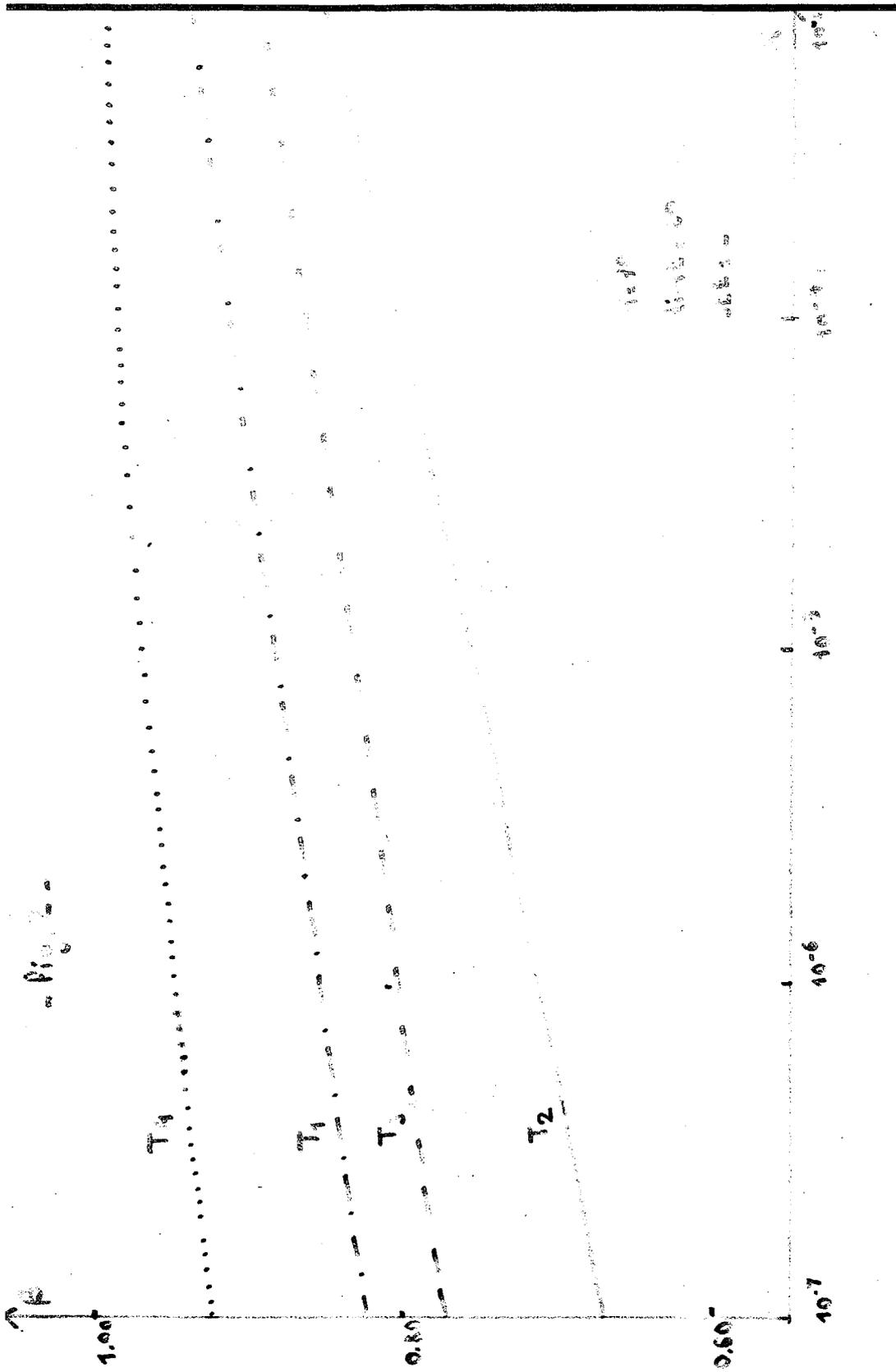


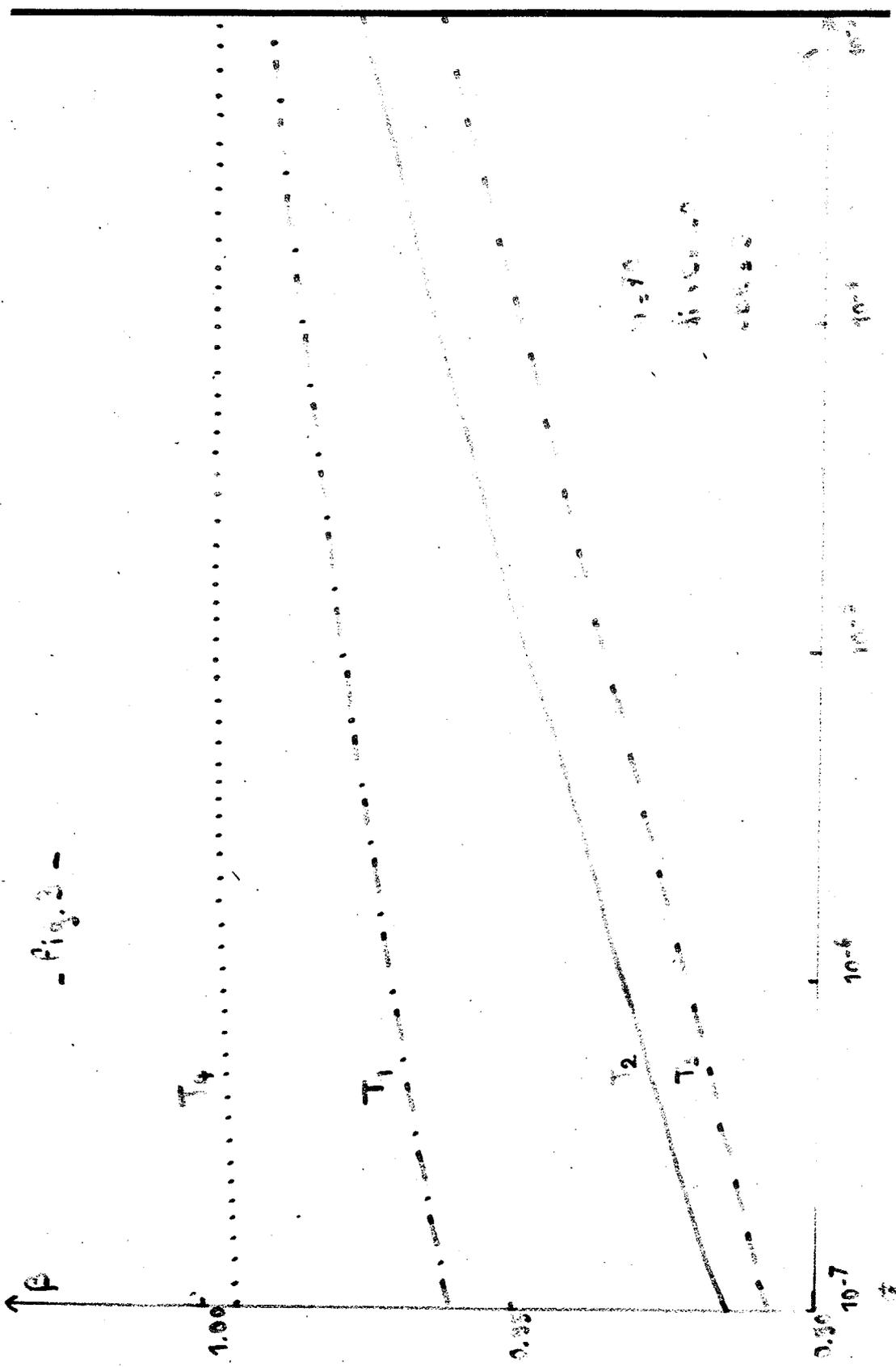
REFERENCES

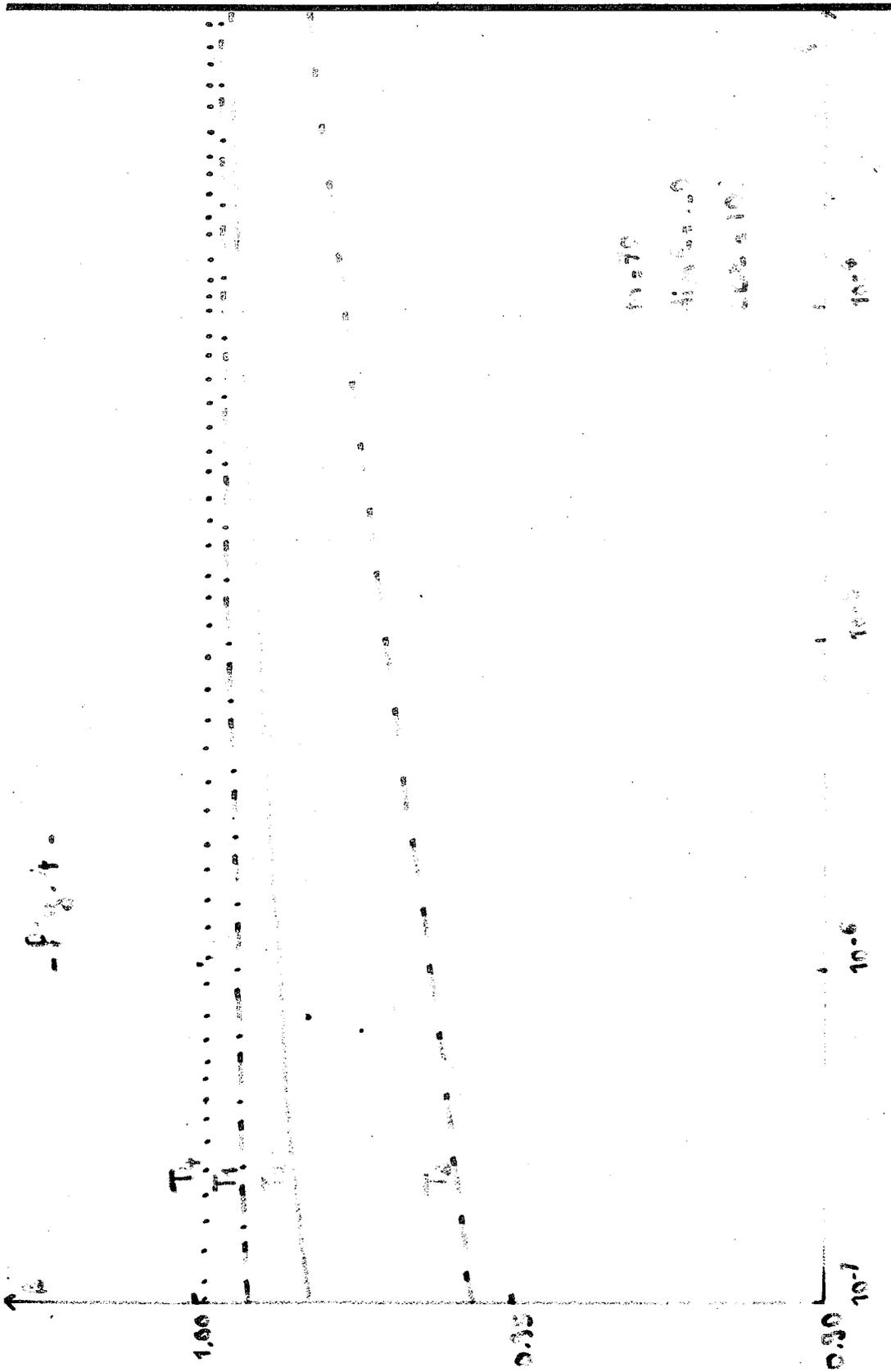
- 1 MANOTTE : "Etude d'échos de sous-marins correspondant à des impulsions F.M. de faible durée".
Colloque GRETSI 1971 (p. 305-322).
- 2 RUTHROFF C.L. : "Rain attenuation and radio path design".
Bell System Technical Journal, Vol. 49, (1), January 1970, (p. 121-135).
- 3 HELSTROM C.W. : "Statistical theory of signal detection".
Second Edition. Pergamon Press, London - New York-1968.
- 4 VAN TREES H.L. : "Detection, Estimation and Modulation Theory".
Part. I : Detection, Estimation and Linear Modulation Theory.
John Wiley and Sons, New York, 1968.
- 5 KAILATH T. : "Correlation Detection of Signals Perturbed by a Random Channel".
IRE Transactions on Information Theory, Vol. IT-6, June 1960 (p. 361-366).
- 6 WAINSTEIN L.A. and ZUBAKOV V.D. : "Extraction of signals from noise".
Prentice Hall, Englewood Cliffs, 1962 (Chap. 5, sec. 35).
- 7 MACCHI Odile : "Détection optimale d'un signal de temps d'arrivée inconnu et estimation de ce temps".
Annales des Télécommunications, t. 25, n° 7-8, Juillet-Août 1970 (p. 312-318).
- 8 FERGUSON T.S. : "Mathematical Statistics, a decision theoretical approach".
Academic Press, New York and London, 1967 (Chap. 4).
- 9 FERGUSON T.S. : -d°- (paragraphe 5-9).
- 10 SOURROUILLE L. : "Détection dans un bruit d'un signal perturbé par un filtrage linéaire aléatoire".
Thèse de 3ème cycle, Orsay, Février 1973.

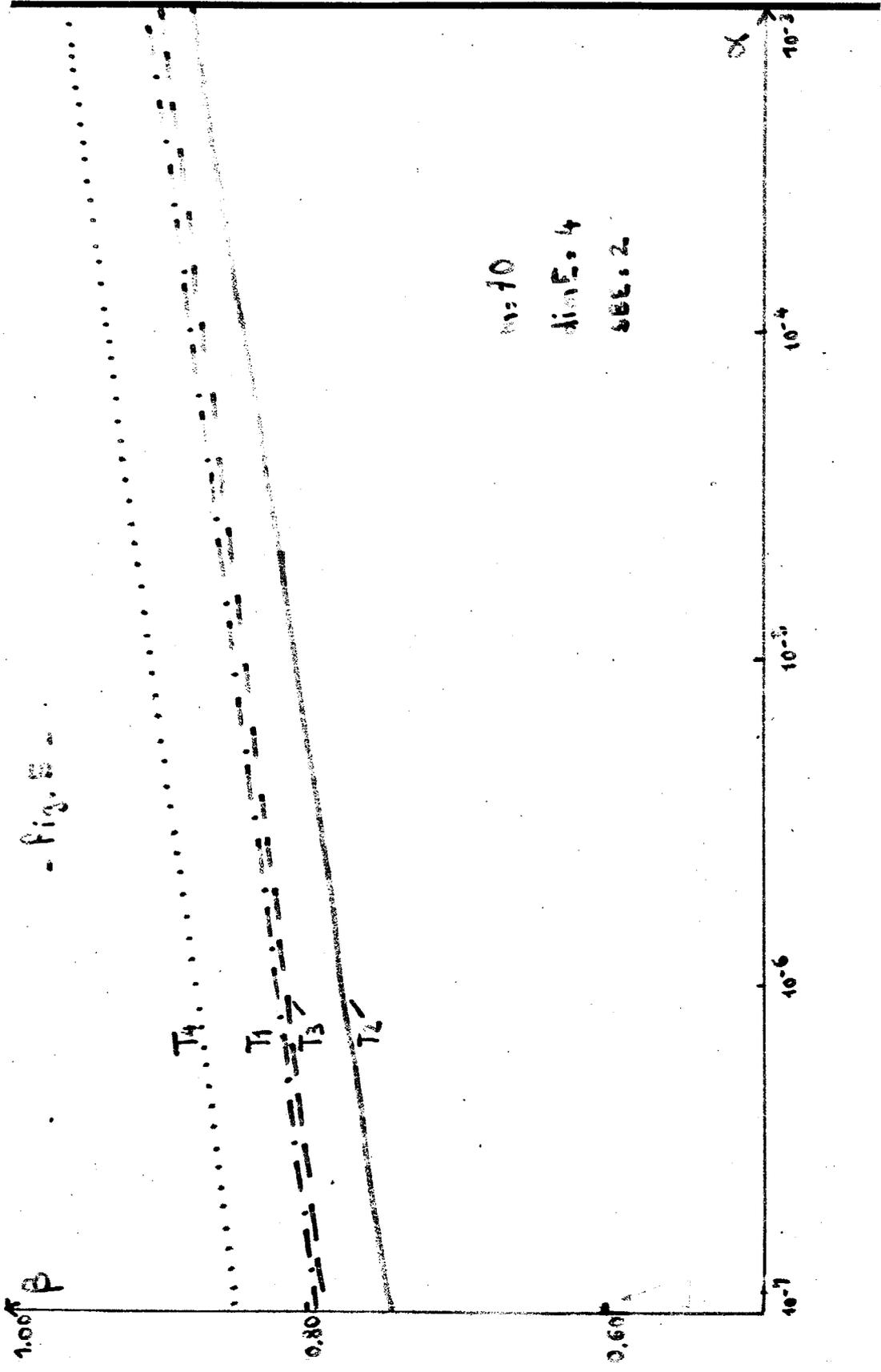


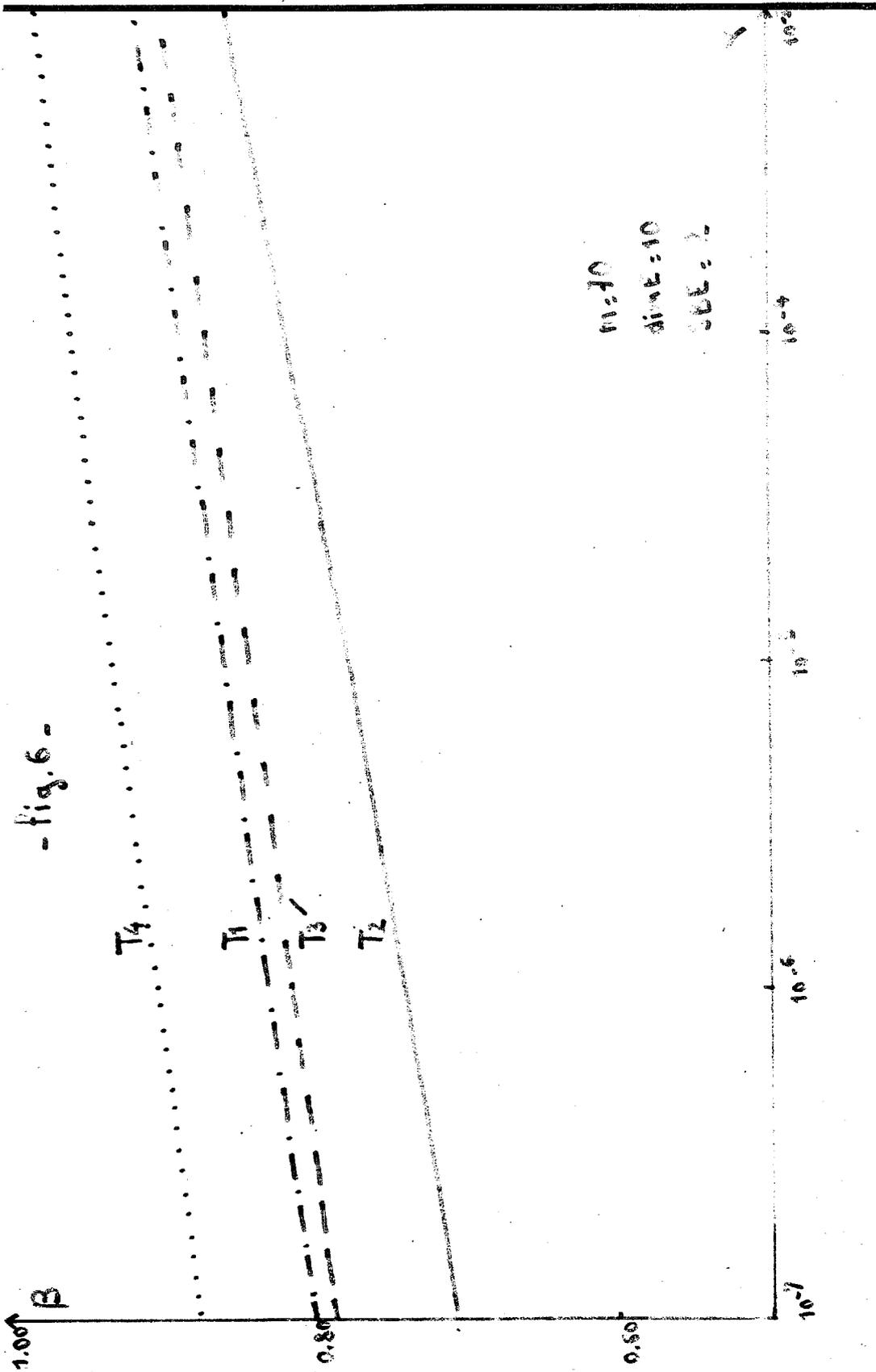
1-70
413682
16541

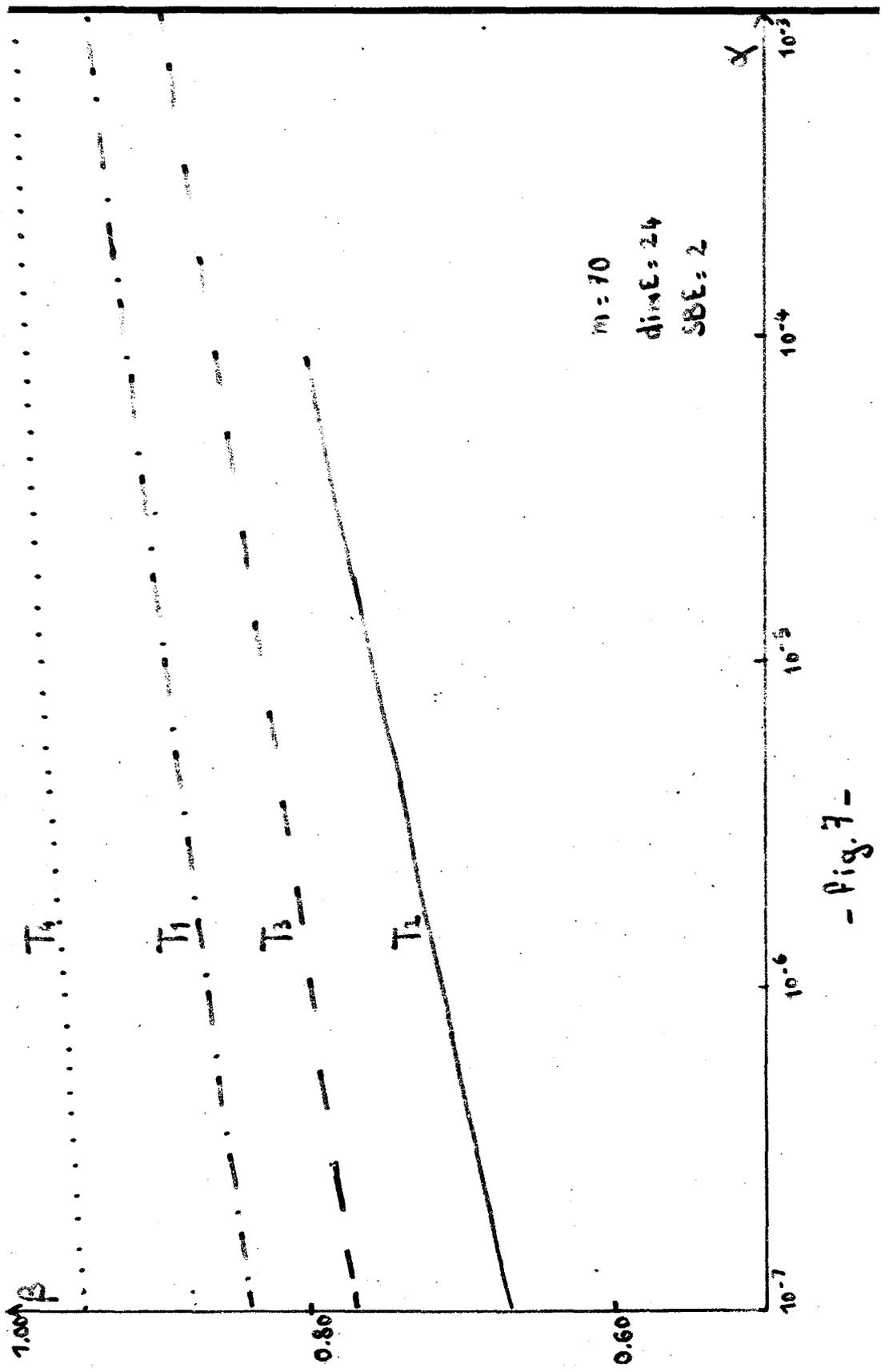












$m = 70$
 $dim E = 24$
 $SBE = 2$

- Fig. 7 -