A Proposed Method for Fast Determination of Plasma Parameters

B.J. Braams[1] and K. Lackner

IPP 1/228 September 1984

MAX-PLANCK-INSTITUT FÜR PLASMAPHYSIK

8046 GARCHING BEI MÜNCHEN

MAX-PLANCK-INSTITUT FÜR PLASMAPHYSIK

GARCHING BEI MÜNCHEN

A Proposed Method for Fast Determination of Plasma Parameters

B.J. Braams[1] and K. Lackner

IPP 1/228 September 1984

[1] On attachment from F.O.M. Instituut voor Plasmaphysica, Postbus 1207, 3430 BE Nieuwegein, The Netherlands.

Die nachstehende Arbeit wurde im Rahmen des Vertrages zwischen dem Max-Planck-Institut für Plasmaphysik und der Europäischen Atomgemeinschaft über die Zusammenarbeit auf dem Gebiete der Plasmaphysik durchgeführt.

IPP 1/228
A Proposed Method for Fast Determination of Plasma Parameters

B.J. BRAAMS[1] AND K. LACKNER

September 1984

ABSTRACT

The method of function parametrization, developed and applied by H. Wind for fast data evaluation in high energy physics, is presented in the context of controlled fusion research. This method relies on statistical analysis of a data base of simulated experiments in order to obtain a functional representation for the intrinsic physical parameters of a system in terms of the values of the measurements. Some variations on Wind's original procedure are suggested. A specific application for tokamak experiments would be the determination of certain global parameters of the plasma, characterizing the current profile, shape of the cross-section, plasma pressure, and the internal inductance. The relevant measurements for this application include values of the poloidal field and flux external to the plasma, and a diamagnetic measurement. These may be combined with other diagnostics, such as electron-cyclotron emission and laser interferometry, in order to obtain also density and temperature profiles. There appears to be a capability for on-line determination of basic physical parameters, in a millisecond timescale on a minicomputer instead of in seconds on a large mainframe.

[1]On attachment from F.O.M. Instituut voor Plasmaphysica, Postbus 1207, 3430 BE Nieuwegein, The Netherlands

1. INTRODUCTION

The following situation is common in experimental physics: the measurements that are made in the course of an experiment do not immediately correspond to the physical parameters that are to be determined, but seem to require for their interpretation a complicated multidimensional parameter fitting procedure. This is the case with virtually any diagnostic for which the recorded signals are a nonlinear function of two or more unknown parameters. Situations occur where the amount of experimental information that is utilized is not limited by the rate at which measurements can be made, but more by the rate at which the raw diagnostics can be interpreted. Clearly then an acceleration of the data evaluation process is desirable.

An important example from controlled fusion research is the determination of plasma parameters during a discharge. MHD equilibrium of a tokamak plasma, specifically, is governed by a two-dimensional partial differential equation with some unknown parameters to describe the current profile. This profile cannot be measured directly, but may be estimated from measurements of the magnetic field made outside the plasma. To interpret such measurements with the aid of the physical model then requires an optimization process in the parameter space, involving repeated solution of the governing p.d.e. This is an expensive process, and not suitable for on-line monitoring of the evolution of the system on a millisecond timescale. (MHD equilibrium refers to a timescale well below 1 ms).

It was recognized by Wind [1], [2], in the context of momentum determination from spark chamber data, that it may be possible to largely dispense with the physical model, and instead employ some relatively simple functional representation for the relation between experimental data and physical parameters. This representation, which gives the estimated physical parameters as an explicit function of the measurements, is found by analysis of a large data set of simulated experiments. A physical model is required in order to generate the data set, but it need not be invoked to find a good functional representation, nor to analyze subsequently the real measurements.

Wind's method of experimental data evaluation through function parametrization consists of three stages. (1) A code is written to model the experiment and the relevant diagnostics. This code is used to generate a data base of simulated states of the system, in which each state is represented by (at least) the values of the physical parameters and of the associated measurements. (2) This data base is made the object of a statistical analysis, involving techniques for dimension reduction and multiple regression, with the aim to provide a relatively simple function expressing the physical parameters in terms of the measurements. (3) The resulting function is then employed for the fast

interpretation of real measurements.

In addition to supplying a function for the interpretation of 'good' data, the analysis also yields criteria with which to test whether actual measurements are consistent with the model that was used to generate the data base, and it yields a procedure to automatically correct erroneous measurements in a near-optimal manner. A substantial effort may be required to generate and analyse the data base, and for the procedure to be economical it is essential that many measurements will be made with the same experimental setup. In such a case the ultimate gain in diagnostic capabilities can be very large.

Function parametrization has been used successfully on several experiments in high energy physics (see references in [**2**]), but it seems not to have become known outside that field. Our purpose with this report is to present this method with a view to experimental data evaluation in fusion research. In section 2 an outline is given of function parametrization as it has been applied in high energy physics. Section 3 contains a discussion of some variations on the original procedure. In section 4 we discuss possible applications in tokamak physics, in particular the rapid determination of the magnetic configuration.

This report is addressed primarily to plasma physicists, although the general presentation in sections 2 and 3 could be of interest to a wider audience. We attempted to make that presentation reasonably self-contained without assuming familiarity with the methods of multivariate statistical analysis.

2. Function parametrization — Outline of Wind's methods

The function parametrization process as it has been employed in high energy physics will now be outlined in more precise terms.

2.1. Preliminaries. A classical physical system is considered, of which $\mathcal{P}$ denotes a typical state. The system may have any number of degrees of freedom, but interest will be restricted to a (partial) characterization by n intrinsic real parameters, represented collectively by a point $\mathbf{p} \in \mathbf{R}^n$. n_0 of these parameters will be independent. In the experimental situation $\mathbf{p}$ is to be estimated from the readings of m measurements ($m \geq n_0$; in many cases $m \gg n_0$), represented by a point $\mathbf{q} \in \mathbf{R}^m$. It is assumed that $\mathbf{p}$ is completely specified by $\mathcal{P}$, but that $\mathbf{q}$ may be a stochastic function of $\mathcal{P}$, the stochasticity being due to random errors in the measurement process. We will write $\mathbf{p} = \mathbf{p}(\mathcal{P})$ and $\mathbf{q} = \mathbf{q}(\mathcal{P})$.

The aim of the function parametrization is to obtain some relatively simple[1] function $\mathbf{F} : \mathbf{R}^m \to \mathbf{R}^n$, such that for any state $\mathcal{P}$ the associated $\mathbf{p}$ and $\mathbf{q}$ satisfy $\mathbf{p} = \mathbf{F}(\mathbf{q}) + \mathbf{e}$ with a sufficiently small error term $\mathbf{e}$. (In the interpretation of experimental data, the parameters $\mathbf{p}_{est}$ may then be estimated from the measurements $\mathbf{q}_{obs}$ by setting $\mathbf{p}_{est} = \mathbf{F}(\mathbf{q}_{obs})$). $\mathbf{F}$ is to be found by analysis of a data base containing the values of the parameters $\mathbf{p}_\alpha$ and of the measurements $\mathbf{q}_\alpha$ for N simulated states $\mathcal{P}_\alpha$ ($1 \leq \alpha \leq N$). This is a problem of function fitting over scattered data in the high-dimensional space $\mathbf{R}^m$, for which techniques from multivariate statistical analysis are appropriate. To a statistician the $\mathbf{q}_\alpha$ are the 'conditions', the $\mathbf{p}_\alpha$ are the 'responses', and $\mathbf{F}$ is a regression. (The terminology of conditions and responses is very unnatural in the present context, and will not be employed in this paper). For the applications that we have in mind, $m \sim 100$, $n_0 < 10$, and $N \sim 10^4$.

2.2. The data base. In a preparatory stage a code G is employed to generate a data base. This code must be suited to compute possible states of the physical system over the whole of the system's regime, and must also contain a model for the measurements. G will take certain numerically convenient parameters (perhaps the components of $\mathbf{p}$) as input, and produce $\mathbf{p}$ and $\mathbf{q}$ as results. The input parameters are varied in a random or a systematic fashion, and for each successful calculation α the values $\mathbf{p}_\alpha$ and $\mathbf{q}_\alpha$ are saved. As the subsequent automatic analysis will only detect dependencies which are reflected in the data base, one must ensure that every parameter or combination of parameters that can vary in the actual experiment is also varied when generating these data.

[1] based on low-order polynomials in only a few linear combinations of the components of $\mathbf{q}$.

It is advisable to store in the data base 'pure' values of the pseudo-measurements, and not to model at this stage the random errors inherent in the real diagnostics. Some considerations related to measurement errors are given in section 3.3.

It need hardly be stressed that the result of the function parametrization can never be expected to improve upon the physical model that is embodied in G, nor can it be employed to find the values of physical parameters that are not reasonably well reflected in the measurements. The traditional least squares method of data interpretation would actually rely on a code such as G, and attempt in an iterative procedure to optimize the free parameters for minimal deviation between simulated and actual measurements. With function parametrization the aim is to obtain a direct and much simpler connection between measurements and physical parameters, without losing too much accuracy.

2.3. Transformation of the measurements. Since the dimensionality m of the space of the measurements lies between several tens and several hundred in many cases, and since a linear fitting function (of the form $\mathbf{p} = \mathbf{p}_0 + \mathbf{A}\mathbf{q}$) is not expected to suffice, the dimensionality of the space of trial functions with which the parameters will be fitted can be very large. A polynomial model of degree l, for instance, will have $n\binom{m+l}{l}$, or $\sim nm^l/l!$ degrees of freedom. It is neccesary to first reduce the number of independent variables by means of a transformation to a lower-dimensional space.

A second, and also very important, aim for this transformation of variables must be to eliminate or reduce multicollinearity (near linear dependencies) between the data points, and thus to improve the conditioning of the regression problem [**3**, ch. 8]. Multicollinearity is likely to be present whenever $m \gg n_0$; specific causes may be some underlying smoothness in the data, or any explicit physical constraint that connects different measurements.

Wind uses a principal component analysis (PCA) [**4**, ch. 8] for the dimension reduction and transformation to independence, and that method will be outlined here. In section 3.1 the use of canonical correlation analysis (CCA) [**4**, ch. 10] will be suggested as an alternative. The reader is referred also to [**3**, ch. 8] and to [**5**, ch. 43] for a discussion of these transformations.

From the N suitably scaled pseudo-measurements $\mathbf{q}_\alpha$, each of which is a point in $\mathbf{R}^m$, the sample mean, $\bar{\mathbf{q}} := N^{-1}\sum_\alpha \mathbf{q}_\alpha$, and the $m \times m$ sample dispersion matrix,

$$\mathbf{S}_{11} := N^{-1}\sum\nolimits_\alpha (\mathbf{q}_\alpha - \bar{\mathbf{q}})(\mathbf{q}_\alpha - \bar{\mathbf{q}})^T,$$

are calculated. $\mathbf{S}_{11}$ is symmetric and positive semi-definite. An eigenanalysis yields m eigenvalues $\lambda_1^2 \geq \cdots \geq \lambda_m^2 \geq 0$, with corresponding orthonormal eigenvectors

$\mathbf{a}_1, \ldots, \mathbf{a}_m$. Any measurement vector $\mathbf{q}$ may be resolved along these eigenvectors to obtain a set of transformed measurements, $x_i := \mathbf{a}_i \cdot (\mathbf{q} - \bar{\mathbf{q}})$. These transformed measurements are linearly independent within the sample, have zero mean, and variance λ_i.

One of the aims, the reduction of multicollinearity, is thereby achieved, but, if all m components x_i are retained, the dimensionality of the problem is not reduced. The credo of PCA is that most information will be contained in the first few 'principal components' $(x_i)_{1 \leq i \leq s}$, where $s \leq m$, and preferably $s \ll m$. The associated first s eigenvectors $\mathbf{a}_i$ are called the 'significant variables'. The principal component transformation is the mapping $\mathbf{R}^m \to \mathbf{R}^s$ defined by $\mathbf{x} = \mathbf{A}^T \cdot (\mathbf{q} - \bar{\mathbf{q}})$, where $\mathbf{A}$ is the matrix which has columns $\mathbf{a}_i$ $(1 \leq i \leq s)$.

Let us briefly discuss the assumptions underlying the use of PCA. The basic assumption is that those linear[1] combinations of the measurements that display the largest relative variance, are also best suited for the interpretation of the data. This requires that the measurements have been properly scaled; beyond that it implies a requirement that the diagnostics are all relevant to the data interpretation, and that variations in any of the physical parameters are well reflected as variations in the measurements. These assumptions may correspond to the experimentalist's aims; in practice, though, significant information may well be concealed in linear combinations of the measurements that show relatively little variation. A preliminary linear or nonlinear transformation of variables based on physical insight can then be beneficial. PCA is invariant only under orthogonal transformations; it is not invariant under scaling of variables, or under more general linear transformations.

If n_0 physical parameters each make a distinguishable contribution to the measurements, $\mathbf{q}$, then at least the first n_0 eigenvalues of $\mathbf{S}_{11}$ will be non-zero. The subsequent eigenvalues need not all vanish (they would vanish if there were a linear relation between $\mathbf{p}$ and $\mathbf{q}$), but they are expected to decrease to much smaller values. This property can be employed to test actual measurements for consistency with the model that was used to generate the data base: projecting an actual measurement vector on the subspace spanned by the least significant variables, one expects to obtain accordingly small components x_i. Any measurement in which one of the 'insignificant' x_i turns out substantially larger in magnitude than the corresponding variance λ_i is suspect. The same principle can be employed to automatically supply a missing component q_k in an actual measurement; one simply sets q_k to that value for which the sum of squares of

[1] 'affine' would be more precise here. We always understand the adjective 'linear' to allow also a constant term.

the scaled insignificant components, x_i/λ_i, is minimized. An interesting application to pattern recognition is described in [**2a**].

2.4. Transformation in parameter space. A procedure similar to the one outlined above is applied to obtain a transformation in parameter space, $\mathbf{p} \to \mathbf{y}$. This transformation should usually be invertible; the aim is not to reduce the dimensionality, but only to eliminate multicollinearity among the parameters that are to be fitted.

One now defines the sample mean of the parameters, $\bar{\mathbf{p}} := N^{-1} \sum_\alpha \mathbf{p}_\alpha$, and the $n \times n$ sample dispersion matrix,

$$\mathbf{S}_{22} := N^{-1} \sum\nolimits_\alpha (\mathbf{p}_\alpha - \bar{\mathbf{p}})(\mathbf{p}_\alpha - \bar{\mathbf{p}})^T,$$

and performs an eigenanalysis on $\mathbf{S}_{22}$. This yields n eigenvalues, $\mu_1^2 \geq \cdots \geq \mu_n^2 \geq 0$, with orthonormal eigenvectors $\mathbf{b}_1 \ldots \mathbf{b}_n$. The transformed parameters $\mathbf{y}$ are then defined by $\mathbf{y} := \mathbf{B}^T \cdot (\mathbf{p} - \bar{\mathbf{p}})$, where $\mathbf{B}$ is the matrix which has columns $\mathbf{b}_j$ $(1 \leq j \leq n)$.

2.5. Regression analysis. Having defined the transformations $\mathbf{q} \to \mathbf{x}$ and $\mathbf{p} \to \mathbf{y}$ it is next neccesary to face the task of fitting the, in general nonlinear, relation between $\mathbf{x}$ and $\mathbf{y}$. The problem has been simplified by the dimension reduction obtained with the transformation $\mathbf{q} \to \mathbf{x}$, and it is expected to be better conditioned through the elimination of multicollinearity.

It is desired to find for each component y_j a regression, $y_j = f_j(\mathbf{x}) + \epsilon_j$, to fit the data $(\mathbf{x}_\alpha, \mathbf{y}_\alpha)_{1 \leq \alpha \leq N}$. These n regressions may be carried out independent of one another, so from here on in this section the index j is discarded and only a single variable y is considered. Wind suggests to use a polynomial model of the form

$$y = \sum_{\mathbf{k}} c_{\mathbf{k}} \cdot \prod_{i=1}^{s} \phi_{k_i}(x_i/r_i) + \epsilon.$$

The multi-index $\mathbf{k}$ has s components $k_1, \ldots, k_s$ in the nonnegative integers, $(c_{\mathbf{k}})_{\mathbf{k}}$ are the unknown regression coefficients, $(\phi_\ell)_{\ell \geq 0}$ is some family of polynomials, r_i is a suitable scaling factor for the variable x_i, and ϵ is the error term. Wind employs Chebyshev polynomials and sets $r_i := \max_\alpha |x_{\alpha i}|$.

An upper bound on some norm of $\mathbf{k}$ must be supplied in order to make the model finite, but for a reasonable bound the model would still have very many terms. This leads one to employ with the above model some form of subset regression, the objective being to retain in the final expression only the terms which make a significant contribution to the goodness-of-fit. A variety of algorithms exists for deciding which terms to retain and which to discard; see for instance [**3**, ch. 7], [**6**, ch. 6], or [**7**, ch. 12]. Wind employs

a simple form of forward selection; his procedure is described in [**1**] and in [**2b**], and is also available as a published algorithm [**8b**].

At this point we have obtained the desired function parametrization, $\mathbf{p} = \mathbf{F}(\mathbf{q}) + \mathbf{e}$, in the following form:

$$\mathbf{p} = \bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{f}(\mathbf{A}^T \cdot (\mathbf{q} - \bar{\mathbf{q}})) + \mathbf{e}$$

with $\bar{\mathbf{q}}$ and $\mathbf{A}$ defined in section 2.3, $\bar{\mathbf{p}}$ and $\mathbf{B}$ defined in 2.4, and $\mathbf{f}$ defined in this section.

2.6. Use of a priori knowledge. From knowledge of the physics of the system it may be possible to provide some a priori approximation $\mathbf{F}_0$ to $\mathbf{F}$. If $\mathbf{F}_0$ is sufficiently easy to evaluate then instead of fitting $\mathbf{F}$ to $(\mathbf{q}_\alpha, \mathbf{p}_\alpha)_{1 \le \alpha \le N}$ one could fit $\mathbf{F}_1$ to the residuals $\mathbf{p}'_\alpha := \mathbf{p}_\alpha - \mathbf{F}_0(\mathbf{q}_\alpha)$, and define $\mathbf{F} := \mathbf{F}_0 + \mathbf{F}_1$. The trivial case $\mathbf{F}_0 := \mathbf{0}$ is allowed in this context, and this should probably be used if one would only be able to provide a *linear* $\mathbf{F}_0$, because the linear terms in the fit are easily found by the automatic procedure anyway (particularly if CCA is employed for the dimension reduction; cf. section 3.1).

A priori knowledge should certainly be employed to choose any initial transformation of the measurements and the parameters (as already discussed in sect. 2.3). The standard automatic analysis is restricted to a polynomial model for $\mathbf{p}$ in terms of $\mathbf{q}$, and this may well be more effective with one representation than with another.

2.7. Conclusion. This completes our brief presentation of Wind's construction of a function parametrization. The stages in sections 2.3 and 2.4 are fairly straightforward, and the difficulty seems to lie mainly with the polynomial fit of section 2.5. The objective is to obtain small residuals $\|\mathbf{p}_\alpha - \mathbf{F}(\mathbf{q}_\alpha)\|$ for an acceptably small number of terms in the polynomial expansion, and it cannot be guaranteed a priori that this objective will be attained when the procedure is tried on the data base for a new experiment. Certainly the procedure should not be expected to work 'automatically'; physical knowledge must be employed to select a good representation for the measurements and the parameters, and should also be the guide when more complicated models than the polynomial regression of section 2.5 seem to be required.

3. Function parametrization -- Variations

Function parametrization in the form as presented in section 2 has been used successfully in high energy physics. However, 'fitting equations to data' is a rich field, as is exhibited clearly in [**9**], and different methods should be tried in order to obtain optimal results. Some alternative approaches are suggested in this section.

3.1. Selection of the significant variables. In Wind's approach, employing principal component analysis, the significant variables are choosen as the first s eigenvectors of the dispersion matrix, $\mathbf{S}_{11} := N^{-1} \sum_\alpha (\mathbf{q}_\alpha - \bar{\mathbf{q}})(\mathbf{q}_\alpha - \bar{\mathbf{q}})^T$. The underlying assumption is that good experimental design and proper scaling of the components of $\mathbf{q}$ has ensured that those linear combinations which display the largest variance are also the most suited for the interpretation of the measurements. In fact, what is wanted of the significant variables $\mathbf{a}_i$ $(1 \leq i \leq s)$ is that the physical parameters $\mathbf{p}$ can be well represented as a function of the transformed measurements $\mathbf{x} \in \mathbf{R}^s$ with components $x_i := \mathbf{a}_i \cdot (\mathbf{q} - \bar{\mathbf{q}})$, with s much less than m.

This indicates that the most appropriate statistical technique for dimension reduction is not principal component analysis, but canonical correlation analysis. In textbooks on multivariate analysis and regression analysis one finds that both techniques are used in practice, although PCA is much more popular. See e.g. [**3**, 8.5], [**4**, 8.8], [**4**, exercise 10.2.8] [**5**, 43.21], or [**6**, 6.9].

With canonical correlation analysis, two transformations, $\mathbf{x} = \mathbf{A}^T \cdot (\mathbf{q} - \bar{\mathbf{q}})$ and $\mathbf{y} = \mathbf{B}^T \cdot (\mathbf{p} - \bar{\mathbf{p}})$, are obtained together. $\mathbf{A}$ and $\mathbf{B}$ are $m \times s$ and $n \times s$ matrices respectively, with columns $\mathbf{a}_i$ and $\mathbf{b}_i$ $(1 \leq i \leq s)$. $\mathbf{a}_1$ and $\mathbf{b}_1$ are choosen so that $x_1 = \mathbf{a}_1 \cdot (\mathbf{q} - \bar{\mathbf{q}})$ and $y_1 = \mathbf{b}_1 \cdot (\mathbf{p} - \bar{\mathbf{p}})$ have the largest possible linear correlation within the sample $(\mathbf{q}_\alpha, \mathbf{p}_\alpha)_{1 \leq \alpha \leq N}$. y_1 is therefore that linear function of the parameters that is best predicted by a linear function of the measurements, and x_1 is the corresponding linear function of the measurements. Generally, each pair $(\mathbf{a}_i, \mathbf{b}_i)$ is choosen to maximize the correlation between x_i and y_i, subject to $\mathbf{a}_i$ being linearly uncorrelated with $\mathbf{a}_1 \ldots \mathbf{a}_{i-1}$, and $\mathbf{b}_i$ being linearly uncorrelated with $\mathbf{b}_1 \ldots \mathbf{b}_{i-1}$. We refer to [**4**, ch. 10] for the mathematical development.

One would expect CCA to be a better technique for dimension reduction than PCA in those cases where the relation between measurements and physical parameters is approximately linear. CCA selects those linear transformations of the measurements that provide the best prediction for some linear transformation of the physical parameters, and disregards combinations which may show a large variance, but are irrelevant to the data interpretation. On the other hand, if a certain linear combination of the measurements enters only quadratically in the regression, then CCA would not select

this combination, although it may be important, and could be found by PCA. So in any case one does well to consider carefully the underlying physics, and also to cross-check the results of a canonical correlation analysis with the aid of a PCA.

One further remark related to CCA is in order: For the present application it will usually be advisable to employ some form of ridge analysis [**3**, ch. 8], [**4**, ch. 10], as the data may be highly co-linear

3.2. Choice of polynomial model. With reference to the general polynomial model of section 2.5, $y = \sum_{\mathbf{k}} c_{\mathbf{k}} \cdot \prod_{i=1}^{s} \phi_{k_i}(x_i/r_i) + \epsilon$, it may be enquired whether there is any reason to prefer one family of polynomials $(\phi_\ell)_{\ell \geq 0}$ to another, choosing among, say, monomials, Chebyshev, Legendre, and Hermite polynomials.

In view of the fact that the first ℓ elements of any family of interest all span the same space, Wind's preference for Chebyshev polynomials is justified only if one is able to choose the transformed measurements $\mathbf{x}_\alpha$ at an appropriate (Chebyshev) set of points. In that case the polynomial model becomes exactly orthogonal, which is a desirable property with regression analysis. Orthogonality implies that all coefficients $c_{\mathbf{k}}$ can be determined independent of one another, and that the choice of whether or not to retain a particular term has no influence on the coefficients of the other terms in the series.

However, it is only in special cases that one has the freedom to choose the $\mathbf{x}_\alpha$ at a Chebyshev set of points. In the standard case discussed above the $\mathbf{x}_\alpha$ are obtained by a transformation of the $\mathbf{q}_\alpha$ in the data base, and already the $\mathbf{q}_\alpha$ are irregularly distributed. The consideration remains that near-orthogonality is a desirable property, and that the correlations between the separate terms in the regression function should be minimized. For that reason we propose to use Hermite polynomials and to set $r_i := \lambda_i$. Within the sample the x_i are linearly uncorrelated with zero mean and variance λ_i, so it is sound to employ a model which is exactly orthogonal over independent normal variables with the same mean and variance. The Hermite polynomial model has that property.

3.3. Errors in variables. In applications of function parametrization, unlike in the standard textbook regression problem, the 'independent' variables $\mathbf{q}$ are not known precisely, but are measured with random deviations. This situation is commonly referred to as 'errors in variables'. Consideration must be given to the effect of these errors on the accuracy of the constructed function $\mathbf{F}$, and also to ways in which $\mathbf{F}$ may be improved in the light of this effect.

In the first place it is important to be aware of a circumstance under which errors in variables completely invalidate the original construction of the function parameterization. This happens whenever there exists some particular linear combination ξ of the transformed measurements x_i, for which the variation in ξ due to random errors in $\mathbf{q}$

dominates the intrinsic variation that is due to the spread of the physical parameters represented in the data base. A careful application of PCA (proper scaling of variables, and not choosing s too large) or of CCA (proper choice of ridge technique) should prevent this circumstance.

Next we mention an a posteriori technique to correct $\mathbf{F}$ for the bias introduced by errors in variables. This is the bootstrap [**10**], which was suggested to us by H.N. Linssen. For every $(\mathbf{q}_\alpha, \mathbf{p}_\alpha)$ in the data base one constructs a sample of pseudo-measurements $\mathbf{q}_\alpha^{(k)}$, randomly distributed around $\mathbf{q}_\alpha$ according to the model for the errors in the measurements. The errors $\mathbf{e}_\alpha^{(k)} := \mathbf{p}_\alpha - \mathbf{F}(\mathbf{q}_\alpha^{(k)})$ are computed, and their mean $\bar{\mathbf{e}}_\alpha$ and dispersion σ_α are derived. Since $\mathbf{F}$ is in general nonlinear, the errors in $\mathbf{q}$ contribute both to σ and to $\bar{\mathbf{e}}$. The bootstrap technique is then to recompute $\mathbf{F}$ in order to compensate as well as is possible for the bias $\bar{\mathbf{e}}$. For the present application this would be done by recomputing the coefficients $c_\mathbf{k}$ in the polynomial fit of section 2.5.

3.4. Miscellaneous remarks. It appears as though function parametrization can be tried, and may well be successful, using the standard techniques described in this and the previous section, but that the efforts of a competent statistician could be of much help. (The present authors are not statisticians).

One issue to be addressed is the use of robust estimators for the means, $\bar{\mathbf{q}}$ and $\bar{\mathbf{p}}$, the dispersion matrices, $\mathbf{S}_{11}$ and $\mathbf{S}_{22}$, and the correlation matrix, $\mathbf{S}_{12}$ (which was implicitely present in section 3.1). Even if normally distributed random parameter values are employed to generate the data base, the nonlinearities in the physics (without which there would be no problem in the first place) may cause the actual values $\mathbf{p}_\alpha$ and $\mathbf{q}_\alpha$ to depart very much from normality.

The use of robust (but biased) estimators for the regression coefficients $c_\mathbf{k}$ may be attractive in conjunction with the bootstrap technique described in section 3.3.

A third issue is the use of other than polynomial models for the regression, e.g. see [**11**] for a very different approach.

4. Determination of a tokamak plasma configuration — an application to fusion research

4.1. Outline. Determination of position and profile of the plasma column in a tokamak discharge seems a particularly indicated application of function parametrization. Reasons for this are:

- A generally accepted and verified physical model exists (the axisymmetric MHD equilibrium equations [**12**]), making the connection between unknown intrinsic plasma parameters, externally applied fields and magnetic measurements a well defined mathematical problem [**13**]–[**16**];
- Knowledge of the position and the profile of the plasma column is the basis for interpretation of practically all other diagnostics. Their identification has therefore to be made at many instants in time, and requires an efficient algorithm. Ultimate aim is in fact on-line real-time analysis with use of the derived information for feedback plasma control;
- Function parametrization allows to connect information from different types of diagnostics: measurements sensitive to the interior structure of flux surfaces can thus be combined with field and flux measurements external to the plasma.

An initial application of function parametrization would be the determination of only a set of global parameters that describe the magnetic configuration (plasma current, position, shape, pressure, and internal inductance), using only the magnetic signals measured outside the plasma. In a more ambitious setup one would wish to determine in a consistent fashion the current, density, and temperature profiles of the plasma, employing an extended set of diagnostics. The two proposals are elaborated below.

4.2. Equilibrium determination from magnetic signals. As a first step we consider only magnetic signals measured outside the plasma. In particular we consider the situation in JET, where in many points along two adjacent surfaces enclosing the toroidal plasma column the poloidal flux function and the tangential component of the poloidal field are measured, and also the change in the total toroidal flux due to currents in the plasma is known. Fast determination of the plasma surface is particularly important in this experiment in view of the variety of possible plasma cross-sections and of the positional instability introduced by the shaping coils and the presence of the iron core.

The relevant measurements would thus be the signals of the 14 poloidal flux loops, the 18 poloidal field coils and one toroidal flux loop, whereas the physical parameters to be determined would be

I_t (toroidal plasma current)

r_0 (major radius of a suitably defined plasma center)

z_0 (vertical displacement of the plasma center)

a (horizontal half-diameter of plasma column)

b/a (elongation of plasma column)

ϑ (triangularity of plasma column)

β_I (poloidal β)

ℓ_i (internal inductance)

The pseudo-experiments to be carried out in preparation for the statistical analysis consist of equilibrium calculations using some fast code for computing free-boundary equilibria. The free parameters of the code are varied in order to cover the whole operating regime of the JET apparatus, and for each calculated equilibrium the physical parameters and the fictitious flux and field measurements are recorded. The dimensionality of the space to be spanned by these calculations is somewhat reduced by the fact that the problem (formulated in terms of equilibrium calculations) is exactly linear in I_t and that the position of the plasma (r_0, z_0), shape of the column $(b/a, \vartheta)$ and horizontal half-diameter a are not independent: definition of the plasma boundary through contact with the limiter constitutes one relation among these 5 parameters. Nevertheless, 6 truly independent parameters remain to be varied, mandating a Monte Carlo approach to the selection of the combinations for which actual calculations have to be carried out. Equilibrium calculations using a Buneman solver on a 64×64 mesh presently require about 300 ms of Cray-1 CPU time per free boundary equilibrium, so that $\sim 10^4$ pseudo-experiments could be carried out per hr of CPU time if overhead operations are kept sufficiently low.

The significant variables in the 33–dimensional space of measurements would be automatically determined by the analysis of the pseudo-measurements; it is however a priori clear on physical grounds that they will be closely related to the lower order Fourier components of the flux and field measurements. The formal analysis will give the exact weighting functions for the contribution of each signal to each of the significant variables.

For this application it is known that $I_t = \oint \mathbf{B} \cdot \mathbf{dl}$, so the formal analysis should deliver very nearly a linear expression for I_t in terms of the measurements of the tangential component of the magnetic field. The analysis given in [**15**] and [**16**] shows that, for a suitable definition of r_0 and z_0, there is a linear expression for $r_0^2 I_t$ and $z_0 I_t$ in terms of the measurements; the latter parameter combinations are then the ones which one should wish to obtain through function parametrization. (One will probably wish to obtain several different characterizations of the plasma position: by moments of the

current distribution; as the center of the outermost flux surface; and as the position of the magnetic axis. An exact expression is available only for the first characterization).

The same references show that β_I and ℓ_i should be replaced, for the function parametrization, by $\beta_I I_t^2$ and $\ell_i I_t^2$, and that the first approximation will be quadratic in the measured magnetic field values.

4.3. Use of additional signals. The limited degree to which information about the intrinsic plasma parameters and the structure of the interior flux surfaces can be gained from external magnetic measurements suggests an attempt to combine such data with other diagnostics.

In particular it is already envisaged to derive the interior flux surface structure as contours of constant T_e from electron-cyclotron emission and Thomson-scattering measurements. Ideally one would like however to utilize for this simultaneously all information about the equilibrium available, and impose also the restriction that the flux-surface structure has to be consistent with solutions to the MHD equilibrium equations. A possible way to do this could be the following.

To the set of measurements is added some representation of the T_e values as obtained with these additional diagnostics. At the same time we select an appropriate basis of functions, $(t_\gamma)_\gamma$, in the normalized flux coordinate $\tilde{\psi} := (\psi - \psi_{edge})/(\psi_{axis} - \psi_{edge})$, and parameterize the temperature profile, $T_e(\tilde{\psi}) = \sum_\gamma \tau_\gamma t_\gamma$. We then employ function parametrization to express all the previous parameters *and* the new parameters τ_γ in terms of the complete set of measurements. Distinguishing feature of this approach is that we use all information about the flux surface structure available, in order to obtain a profile for the temperature that is also consistent with the MHD equilibrium equations.

An additional set of signals to be considered for inclusion in the function parametrization would be Faraday rotation measurements as carried out on TEXTOR [17]. These measurements will in general be too few in number to allow satisfactory interpretation in isolation, particularly at high β and/or for non-circular cross sections, and function parametrization may be the most suitable means by which to extract the information that is contained in these data.

5. Conclusions

We have outlined the method of function parametrization for experimental data evaluation as it has been used previously, and suggested several variations. Function parametrization may be an effective tool in plasma physics as well as in high energy physics. A particularly important application for tokamak experiments would be the rapid determination of global parameters characterizing the plasma current, position, shape, pressure, and internal inductance.

An investigation along these lines is in progress at IPP.

Acknowledgements

The first author learnt of function parametrization in 1973, when he was a summer student at CERN, working under the guidance of H. Wind (on a different project). Valuable conversations with Dr. Wind, and the hospitality of the CERN organisation, are gratefully acknowledged. Discussions with Prof. R. Doornbos and Dr. H.N. Linssen (T.U. Eindhoven) helped to clarify many issues of the statistical analysis. Discussions with Prof. F. Engelmann and Dr. C.E. Singer helped to improve the presentation.

The work of B.J. Braams was performed under the Euratom–F.O.M. association agreement, with financial support from Z.W.O. and Euratom.

REFERENCES

1. H. Wind, *Function Parametrization*, in "Proceedings of the 1972 CERN Computing and Data Processing School", CERN 72-21, 1972, pp. 53-106.
2. H. Wind, (a) *Principal Component Analysis and its Application to Track Finding*; (b) *Interpolation and Function Representation*, in "Formulae and Methods in Experimental Data Evaluation", Vol. 3, European Physical Society, Geneva, 1984.
3. D.C. Montgomery and E.A. Peck, "Introduction to Linear Regression Analysis", Wiley, New York, 1982.
4. K.V. Mardia, J.T. Kent and J.M. Bibby, "Multivariate Analysis", Academic Press, London, 1979.
5. M.G. Kendall and A. Stuart, "The Advanced Theory of Statistics", (3 Vols.), 4th Ed., Griffin, London, 1977-1983.
6. N.R. Draper and H. Smith, "Applied Regression Analysis", 2nd Ed., Wiley, New York, 1981.
7. G.A.F. Seber, "Linear Regression Analysis", Wiley, New York, 1977.
8. M.J. O'Connel, (a) *Search Program for Significant Variables*, Comput. Phys. Commun. **8** (1974), 49-55; (b) *Multivariate Least Squares Fitting Program Using Modified Gram-Schmidt Transformations*, ibid, 56-69.
9. C. Daniel and F.S. Wood, "Fitting Equations to Data; Computer Analysis of Multifactor Data", 2nd ed., Wiley, New York, 1980.
10. E. Bradley, "The Jackknife, the Bootstrap and Other Resampling Plans", SIAM, Philadelphia, 1982.
11. J.H. Friedman and W. Stuetzle, *Projection Pursuit Regression*, J. Amer. Statist. Assoc. **76** (1981), 817-823.

12a. V.D. Shafranov, *On Magnetohydrodynamical Equilibrium Configurations*, Sov. Phys. JETP **6** (1958), 545-554; J. Exper. Theor. Phys. (USSR) **33** (1957), 710-722.

12b. R. Lüst and A. Schlüter, *Axialsymmetrische Magnetohydrodynamische Gleichgewichtskonfigurationen*, Z. Naturforsch. **12a** (1957), 850-854.

12c. H. Grad and H. Rubin, *Hydromagnetic Equilibria and Force-Free Fields*, in "Proceedings of the Second United Nations International Conference on the Peaceful Uses of Atomic Energy", Geneva, 1958.

13. K. Lackner, *Computation of Ideal MHD Equilibria*, Comput. Phys. Commun. **12** (1976), 33-44.
14. T.S. Wang and F.J. Helton, *Algorithms Developed to Obtain Particular Numerical Solutions to the Ideal MHD Equilibrium Problem*, Comput. Phys. Commun **24** (1981), 255-261.
15. L.E. Zakharov and V.D. Shafranov, *Equilibrium of a Toroidal Plasma with Noncircular Cross Section*, Sov. Phys. Tech. Phys **18** (1973), 151-156; Zh. Tekh. Fiz. (USSR) **43** (1973), 225-234.
16. B.J. Braams, *On the Interpretation of Tokamak Magnetic Diagnostics*, Report IPP 1/235. (In preparation).
17. H. Soltwisch, *Measurement of Electron and Current Density Profiles in Textor by Means of a Far-Infrared Interferometer/Polarimeter*, Paper C01, 11th European Conference on Controlled Fusion and Plasma Physics, Aachen, 1983, Europhysics Conference Abstracts **7D** (1983), 123-126.