An Original Approach to Mode Converter Optimum Design

Eric Lunéville, Jean-Michel Krieg, and Eric Giguet

***Abstract*—An original method of shape optimization has been developed to improve high-power high-frequency transmission-lines' performances. This method is based on the coupling coefficients equations, for which general expressions are given. It makes use of numerical methods such as steepest descent associated to adjoint state technique. The results obtained on several types of components demonstrate the pertinence of this method.**

***Index Terms*— Coupled mode analysis, optimization methods, transmission lines.**

I. Introduction

HIGH-FREQUENCY high-level RF power is necessary for electron cyclotron applications (plasma heating and diagnostics) in thermonuclear fusion experiments. To generate such a wave, gyrotron oscillators (tubes which can deliver powers greater than 1-MW continuous wave (CW) at frequencies above 100 GHz) are the most commonly used sources. The power then propagates to the application (tokamak) through a transmission line, which must be oversized to avoid breakdowns and to minimize ohmic losses. The first goal of the line is to transform the mode generated in the gyrotron cavity, usually a high-order mode ($TE_{6,4}$, $TE_{15,2}$, $TE_{22,6}$, etc.) to obtain high-power levels, into another mode more suitable for the application, like the HE_{11} mode propagating in a corrugated waveguide and having the advantages of low ohmic losses and of a quasi-Gaussian radiation pattern. This can be done through a set of passive components called mode converters, which are actually waveguides deformed according to specific laws. Among the other components of a transmission line, one can list tapers, allowing a change of the line radius (e.g., to optimize the conversion from one mode to another) and bends used to change the wave direction of propagation. All these components must satisfy two main criteria: they must be as efficient as possible (more than 95%), for the useful power to be as high as possible, and they must be as short as possible to reduce the dimensions of the set "gyrotron-transmission line."

The most commonly used components are: 1) smooth tapers; 2) sinusoidal converters (i.e., converters where the radius is a sinusoidal function of the longitudinal axis) which can transform a TX_{mp} mode into a $TX'_{m'p'}$ one (where X and X' represent either a TE or TM mode); 3) helical converters (the radius of which not only depends on the longitudinal axis, but also on the asimuthal angle) converting TX_{mp} modes into $TX'_{m'p'}$ ones; and 4) serpentine converters, where the radius is constant, but the longitudinal axis is curved, which allow the same kind of conversion as the previous ones. The amplitudes of the modes propagating in such structures are solutions of a set of first-order differential equations. A suitable efficiency is obtained by varying various parameters, like the slope of the components (tapers), their length, and the amplitude and period of their deformations (converters), to keep the level of the spurious modes as low as possible [1]–[7].

Nevertheless, it is not always possible when using such techniques to find the optimal point in terms of efficiency of length. We have developed a method based on shape optimization techniques to increase the efficiency and reduce the length of transmission lines—and this in an automatic way.

The Section II of this paper deals with the physical model of mode conversion, leading to the resolution of a set of differential equations. The material of this section is partly available in different papers [8]–[11]. We found it necessary to review the final results in the two cases of radius and curvature longitudinal deformations.

In the Section III, we detail the different techniques used to optimize the shape with regard to the conversion efficiency. The main ideas are: 1) cubic spline representation of the shape; 2) conservative Crank–Nicolson scheme to solve the differential system; 3) cost function taking into account the conversion efficiency and other parameters; and 4) adjoint state equations for fast computation of the gradient used in a steepest descent algorithm.

Section IV summarizes the main experimental results obtained on several tested components.

II. Wall Deformations–Coupling Equations

The field components $(\mathbf{E}, \mathbf{H})$ of an electromagnetic wave propagating in a circular waveguide can be written as

$$\mathbf{E}(r, z, \theta, t) = \sum_j V_j(z, t)\hat{e}_j(r, \theta) + \sum_n V_n(z, t)\hat{e}_n(r, \theta) + \sum_n \mathbf{E}_{zn}(r, z, \theta, t) \quad (1)$$

$$\mathbf{H}(r, z, \theta, t) = \sum_j I_j(z, t)\hat{h}_j(r, \theta) + \sum_n I_n(z, t)\hat{h}_n(r, \theta) + \sum_j \mathbf{H}_{zj}(r, z, \theta, t). \quad (2)$$

Manuscript received October 21, 1993; revised October 9, 1997.

E. Lunéville is with S.M.P., URA 853 CNRS, 91120 Palaiseau, France.

J.-M. Krieg was with Thomson Tubes Electroniques, 78141 Vélizy-Villacoublay, France. He is now with the Observatoire de Paris, 75014 Paris, France.

E. Giguet is with Thomson Tubes Electroniques, 78141 Vélizy-Villacoublay, France.

Publisher Item Identifier S 0018-9480(98)00619-X.

The particular waves (represented by the index j) such as $E_2 = 0$ are called TE modes, while those (index n) such as $H_z = 0$ are called TM modes. In the previous equations, the scalar functions V and I are assumed to vary as

$$V(z,t) = Ve^{j(\omega t - k_z z)} \quad I(z,t) = Ie^{j(\omega t - k_z z)}$$

and the vectors $(\hat{e}_j, \hat{h}_j, \hat{e}_n, \hat{h}_n)$ to be normalized on the cross-section S by $\int_S \hat{e}_\lambda, \hat{e}_\eta^* \, dS = \int_S \hat{h}_\lambda \hat{h}_\eta^* \, dS = \int_S (\hat{e}_\lambda \times \hat{h}_\eta^*)\hat{e}_z \, dS = \delta_{\lambda\eta}$ if the modes λ and η represent both TE or both TM modes, these integrals being null otherwise.

The actual values of these vectors can be determined by solving Maxwell's equations $\nabla \times \mathbf{H} = j\omega\epsilon\mathbf{E}$ and $\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H}$ in the case where the wave propagating in the guide is a pure mode.

1) TE Mode: Let us introduce a scalar function $\Psi(r, \theta)$ defined by: $\hat{h}(r, \theta) = -\hat{\nabla}\Psi$; $\hat{e}(r, \theta) = \hat{h} \times \hat{e}_z$. The condition of normality of the electric field on the wall leads to $\frac{\partial \Psi}{\partial r} = 0$ for $r = a$ (where a is the radius of the waveguide). Using the normalization conditions, one can finally write

$$\Psi(r, \theta) = \frac{J_m\left(\mu_{mp}\frac{r}{a}\right)}{\sqrt{\pi} J_m(\mu_{mp})\sqrt{\mu_{mp}^2 - m^2}} \cos m\theta$$

where μ_{mp} is the pth root of J'_m.

2) TM Mode: We introduce the function $\chi(r, \theta)$ defined by: $\hat{e}(r, \theta) = -\hat{\nabla}\chi$; $\hat{h}(r, \theta) = \hat{e}_z \times \hat{e}$. The value of the electric field on the wall and the normalization conditions lead to

$$\chi(r, \theta) = \frac{J_m\left(\nu_{mp}\frac{r}{a}\right)}{\sqrt{\pi} J_m(\nu_{mp})} \sin m\theta$$

where ν_{mp} is the pth root of J_m.

As long as the wavelength is purely cylindrical, the decomposition given by (1) and (2) remains the same. Let us now consider a slightly deformed waveguide, i.e., such as its radius varies with z and θ, or its axis z is curved (even if the radius is constant). Theoretically, because of the change in the normal to the waveguide, the previously given field values are not valid. Nevertheless, if the deformations remain smooth enough, one can consider that the wave propagating can still be represented as a set of TE and TM modes even if the composition varies while it propagates, that is to say that the different modes can be coupled to each other. In this section the coupling equations which must be taken into account in two particular cases, variation of the radius and of the axis curvature will be summarized.

A. Variation of the Radius a

The coupling equations have first been developed in the particular case where the radius depends only on the longitudinal coordinate z. As stated in [8] and [9], in the case of small variations they can be written as

$$\begin{aligned}\frac{dA_i^+}{dz} &= -jk_{zi}A_i^+ + \sum_j C_{ij}^+ A_j^+ + \sum_j C_{ij}^- A_j^- \\ &\quad + \sum_n C_{\text{in}}^+ A_n^+ + \sum_n C_{\text{in}}^- A_n^- \\ \frac{dA_i^-}{dz} &= jk_{zi}A_i^- + \sum_j C_{ij}'^+ A_j^+ + \sum_j C_{ij}'^- A_j^- \\ &\quad + \sum_n C_{\text{in}}'^+ A_n^+ + \sum_n C_{\text{in}}'^- A_n^- \\ \frac{dA_t^+}{dz} &= -jk_{zt}A_t^+ + \sum_j C_{tj}^+ A_j^+ + \sum_j C_{tj}^- A_j^- \\ &\quad + \sum_n C_{tn}^+ A_n^+ + \sum_n C_{tn}^- A_n^- \\ \frac{dA_t^-}{dz} &= jk_{zt}A_t^- + \sum_j C_{tj}'^+ A_j^+ + \sum_j C_{ij}'^- A_j^- \\ &\quad + \sum_n C_{tn}'^+ A_n^+ + \sum_n C_{tn}'^- A_n^-.\end{aligned}$$

In this set of first-order differential equations, $A^{+(-)}$ represents the amplitude of the forward (backward) component of the different modes. These amplitudes are derived from V and I by the relations

$$V = \sqrt{Z}(A^+ + A^-) \quad I_i = \frac{1}{\sqrt{Z}}(A^+ - A^-)$$

with $Z = \frac{\omega\mu_0}{k_z}$ for TE modes and $Z = \frac{k_z}{\omega\epsilon_0}$ for TM modes.

The indexes i and j represent TE modes, t and n TM modes. The coupling coefficient between two modes is nonnull if, and only if, they have the same azimuthal index. In that occurrence, and depending on the type of modes, the coefficients are given by

$$\begin{aligned}C_{ii}^+ &= C_{ii}'^- = C_{tt}^+ = C_{tt}'^- = 0 \\ C_{ij}^+ &= C_{ij}'^- = -\frac{1}{2}\left[F_{ji}\sqrt{\frac{k_{zj}}{k_{zi}}} - F_{ij}\sqrt{\frac{k_{zi}}{k_{zj}}}\right] \\ C_{ii}^- &= C_{ii}'^+ = -\frac{1}{2Z_i}\frac{dZ_i}{dz} \\ C_{ij}^- &= C_{ij}'^+ = \frac{1}{2}\left[F_{ji}\sqrt{\frac{k_{zj}}{k_{zi}}} + F_{ij}\sqrt{\frac{k_{zi}}{k_{zj}}}\right] \\ C_{\text{in}}^+ &= C_{\text{in}}'^- = \frac{1}{2}\frac{\omega}{c}\frac{G_{\text{in}}}{\sqrt{k_{zi}k_{zn}}} C_{\text{in}}^- = C_{\text{in}}'^+ = -\frac{1}{2}\frac{\omega}{c}\frac{G_{\text{in}}}{\sqrt{k_{zi}k_{zt}}} \\ C_{tj}^+ &= C_{tj}^- = C_{tj}'^+ = C_{tj}'^- = \frac{1}{2}\frac{\omega}{c}\frac{F_{tj}}{\sqrt{k_{zj}k_{zt}}} \\ C_{tn}^+ &= C_{tn}'^- = -\frac{1}{2}\left[F_{nt}\sqrt{\frac{k_{zt}}{k_{zn}}} - F_{tn}\sqrt{\frac{k_{zn}}{k_{zt}}}\right] \\ C_{tt}^- &= C_{tt}'^+ = -\frac{1}{2Z_t}\frac{dZ_t}{dz} \\ C_{tn}^- &= C_{tn}'^+ = \frac{1}{2}\left[F_{nt}\sqrt{\frac{k_{zt}}{k_{zn}}} + F_{tn}\sqrt{\frac{k_{zn}}{k_{zt}}}\right]\end{aligned}$$

with α azimuthal index of the modes:

$$\begin{aligned}F_{ii} &= -G_{ii} = -\frac{1}{a}\frac{da}{dz}\frac{\alpha^2}{\mu_i^2 - \alpha^2} \\ F_{ij} &= -G_{ij} = \frac{1}{a}\frac{da}{dz}\frac{2\mu_j^2}{\mu_j^2 - \mu_i^2}\sqrt{\frac{\mu_i^2 - \alpha^2}{\mu_j^2 - \alpha^2}} \\ \mu_{i(j)} &= \mu_{i(j)mp}\,(p\text{th root of } j'_m)\end{aligned}$$

$$G_{\mathbf{in}} = \frac{1}{a}\frac{da}{dz}\frac{2\sigma}{\sqrt{\mu_j^2-\alpha^2}}$$
$$F_{tj} = -\frac{1}{a}\frac{da}{dz}\frac{2\alpha}{\sqrt{\mu_j^2-\alpha^2}}$$
$$v_{\nu(t)} = \nu_{n(t)mp}\,(p\text{th root of } J_m)$$
$$F_{tt} = -G_{tt} = -\frac{1}{a}\frac{da}{dz}$$
$$F_{tn} = -G_{tn} = -\frac{1}{a}\frac{da}{dz}\frac{2\mu_t^2}{\mu_t^2-\mu_n^2}.$$

Some applications require conversion of one mode into another one with a different azimuthal index. This can be done through a component with radial variations in both z and θ. The most commonly used deformation is given by

$$a(z,\theta) = a_0[1+\epsilon\cos(\beta z - l\theta)]$$

where a_0 is the mean radius and l must be equal in absolute value to the difference between the two azimuthal indexes. The coupling coefficients to be used in this case have been determined, particularly by Mourier [12].

B. Serpentine Converter–Morgan Coefficients

Another way to couple modes with different azimuthal indexes consists in the use of a so-called "serpentine" converter. In such a component, the radius remains constant, but the axis of propagation is deformed. The coupling equations to be taken into account in this case have been developed by Morgan [11].

The electromagnetic field of a wave propagating in such a guide can be calculated using the coordinates system (r, θ, w), where w is the distance measured along the axis and (r, θ) the polar coordinates in a plane perpendicular to w and the origin of which is on this axis. In the following, the radius of curvature of w will be noted b.

Developing Maxwell's equations in this coordinate system finally leads [in the particular case where the curvature is small (i.e., $\frac{r}{b}\cos\theta \le \frac{a}{b} \ll 1$)] to the following first-order coupling equations (i and j still represent TE mode, t and n TM modes):

$$\frac{\partial V_i}{\partial w} = \sum_j Z_{ij}I_j + \sum_n Z_{\mathbf{in}}I_n$$
$$\frac{\partial I_i}{\partial w} = \sum_j Y_{ij}V_j + \sum_n Y_{\mathbf{in}}V_n$$
$$\frac{\partial V_t}{\partial w} = \sum_j Z_{tj}I_j + \sum_n Z_{tn}I_n$$
$$\frac{\partial I_t}{\partial w} = \sum_j Y_{tj}V_j + \sum_n Y_{tn}V_n.$$

The coupling coefficients are given by the following:

$$Z_{ij} = j\omega\mu_0(\Theta_{ij}-\delta_{ij})$$
$$Y_{ij} = j\omega\epsilon_0(-\delta_{ij}+\Theta_{ij}) - \frac{K_{ci}K_{cj}}{j\omega\mu_0}(\delta_{ij}+e_{ij})$$
$$K_{ci(j)} = \frac{\mu_{i(j)}}{a}$$
$$Z_{\mathbf{in}} = j\omega\mu_0\Theta_{\mathbf{in}}$$
$$Y_{\mathbf{in}} = j\omega\epsilon_0\Theta_{\mathbf{in}}$$
$$Z_{tj} = j\omega\mu_0\Theta_{tj}$$
$$Y_{tj} = j\omega\epsilon_0\Theta_{tj}$$
$$Z_{tn} = j\omega\mu_0(-\delta_{tn}+\Theta_{tn}) - \frac{K_{cn}K_{ct}}{j\omega\epsilon_0}(\delta_{tn}+e_{tn})$$
$$Y_{tn} = j\omega\varepsilon_0(\Theta_{tn}-\delta_{tn})$$
$$K_{cn(t)} = \frac{v_{n(t)}}{a}.$$

Replacing the indexes i, j, t, and n, respectively, by the double indexes ie, sv, tu, and mn, one finds the following:

$$\Theta_{ij} = \Theta_{ie,\,sv} = -\alpha_{ij}\frac{a}{b}\frac{2\mu_{ie}^2\mu_{sv}^2 - is(\mu_{ie}^2+\mu_{sv}^2)}{(\mu_{sv}^2-\mu_{ie}^2)^2\sqrt{\mu_{sv}^2-s^2}\sqrt{\mu_{ie}^2-i^2}}$$
$$e_{ij} = e_{ie,\,sv} = \alpha_{ij}\frac{a}{b}\mu_{ie}\mu_{sv}\frac{\mu_{ie}^2+\mu_{sv}^2-2s}{(\mu_{sv}^2-\mu_{ie}^2)^2\sqrt{\mu_{ie}^2-i^2}\sqrt{\mu_{sv}^2-s^2}}$$
$$\Theta_{\mathbf{in}} = \Theta_{ie,\,mn} = -\frac{a}{b}\frac{1}{\sqrt{2}\mu_{ie}}, \quad \text{if } i=0,\ m=1,\ \text{and } e=n$$
$$= 0, \quad \text{if } i=0,\ m=1,\ \text{and } e\neq n$$
$$= -\alpha_{\mathbf{in}}\frac{a}{b}\frac{i}{(\mu_{ie}^2-v_{mn}^2)\sqrt{\mu_{ie}^2-i^2}}, \quad \text{if } i\neq 0$$
$$\Theta_{tj} = \Theta_{tu,\,sv} = -\frac{a}{b}\frac{1}{\sqrt{2}\mu_{sv}}, \quad \text{if } t=1,\ s=0,\ \text{and } v=u$$
$$= 0, \quad \text{if } t=1,\ s=0,\ \text{and } v\neq u$$
$$= -\alpha_{tj}\frac{a}{b}\frac{s}{(\mu_{sv}^2-v_{tu}^2)\sqrt{\mu_{sv}^2-s^2}}, \quad \text{if } t\neq 1$$
$$\Theta_{tn} = \Theta_{tu,\,mn} = -\alpha_{tn}\frac{a}{b}\frac{(v_{mn}^2+v_{tu}^2)}{(v_{mn}^2-v_{tu}^2)^2}$$
$$e_{tn} = e_{tu,\,mn} = 2\alpha_{tn}\frac{a}{b}\frac{v_{tu}v_{mn}}{(v_{mn}^2-v_{tu}^2)^2}$$

where

$$\alpha_{ij} = \alpha_{ie,\,sv} = \sqrt{2}, \quad \text{if } i=0 \text{ and } s=1 \text{ or } i=1 \text{ and } s=0$$
$$= 1, \quad \text{if } i\neq 0,\ \text{and } s=i+1 \text{ or } i\neq 1 \text{ and } s=i-1$$
$$= 0, \quad \text{otherwise}$$
$$\alpha_{\mathbf{in}} = \alpha_{ie,\,mn} = 1, \quad \text{if } m=i+1 \text{ or} i\neq 1 \text{ and } m=i-1$$
$$= 0, \quad \text{otherwise}$$
$$\alpha_{tj} = \alpha_{tu,\,sv} = 1, \quad \text{if } s=t-1 \text{ or } t\neq 0 \text{ and } s=t+1$$
$$= 0, \quad \text{otherwise}$$
$$\alpha_{tn} = \alpha_{tu,\,mn} = 1, \quad t\neq 0,\ \text{and } m=t+1\text{or } t\neq 1 \text{ and } m=t-1$$
$$= 0, \quad \text{otherwise.}$$

Using these coupling equations, one can fine the equations in A^+ and A^- using the relations $V=\sqrt{Z}(A^+ + A^-)$ and $I = \frac{1}{\sqrt{Z}}(A^+ - A^-)$.

III. Optimum Design of Waveguides

A. A Class of Optimum Design Problem

As we have seen in the previous section, the mode-coupling model of electromagnetic waves in axisymetric guides or "serpentine" guides leads to a one-dimensional differential system that we state in general form as follows:

$$\begin{cases} \frac{d}{dx}\mathbf{A}(x) = \mathbf{C}(h, x)\mathbf{A}(x), & x \in]0, L[\\ \mathbf{A}^+(0) = \mathbf{A}_0 \end{cases} \tag{3}$$

$$\mathbf{A}^-(L) = 0 \tag{4}$$

where L is the length of the guide, x is the propagation axis

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}^+ \\ \mathbf{A}^- \end{pmatrix}$$

$\mathbf{A}^+$ (respectively, $\mathbf{A}^-$) being the complex vector of the amplitudes of the forward modes (respectively, backward modes)

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}^{++} & \mathbf{C}^{+-} \\ \mathbf{C}^{-+} & \mathbf{C}^{--} \end{bmatrix}$$

$\mathbf{C}^{++}$, $\mathbf{C}^{--}$, $\mathbf{C}^{-+}$, $\mathbf{C}^{+-}$ are, respectively, the matrices of coupling coefficients between forward and backward waves, the real function $h(x)$ is the design parameter of the guide (for instance, h is the radius of the section of an axisymetric guide or the curvature of the axis of a serpentine guide), $\mathbf{A}_0$ is the input vector mode amplitudes.

The output condition (4) means that backward modes do not come from right infinity. For the sake of simplicity, in this paper, we neglect the effects of backward modes. Therefore, we assume $\mathbf{A} = \mathbf{A}^+$, $\mathbf{C} = \mathbf{C}^{++}$ and we deal with the model (3) in the following.

Knowing the input mode $\mathbf{A}_0$, we want to generate at $x = L$ only one given mode (called m). As the matrix $\mathbf{C}$ is anti-Hermitian ($\mathbf{C}^* = -\mathbf{C}$, where $\mathbf{C}^*$ represents the conjugated matrix of $\mathbf{C}$), the system (3) is conservative, i.e.,

$$||\mathbf{A}(x)|| = ||\mathbf{A}_0||, \qquad \text{for every } x \in [0, L]$$

where $||\ ||$ is the norm on the complex space.

Let us now introduce the maximization problem

$$\max_{h \in \mathcal{R}} J(h) \tag{5}$$

where $J(h) = |A_m(L)|^2$ is the cost function, A_m represents the mth component of the vector $\mathbf{A}$ and $\mathcal{R}$ is the set of constraints fulfilled by the design function h.

It is straightforward to interpret this problem in terms of a maximization of energy efficiency, and the best one is reached when

$$|A_m(L)|^2 = ||\mathbf{A}_0||^2.$$

We take two kinds of constraints into account. First of all, structural constraints are expressed as follows:

$$\begin{cases} h(0) = h_0, & h(L) = h_L \\ h'(0) = 0, & h'(L) = 0. \end{cases} \tag{6}$$

These constraints specify geometric properties of guides (for instance, we have $h_0 = h_L$ for axisymetric converters, $h_0 \neq h_L$ for tapers, and $h_0 = h_L = 0$ for serpentine waveguides) and ensure that two successive guides will fit together. The treatment of these constraints will be detailed in the next section.

Secondly, we consider some quality constraints:

on the mode exchange energy:

$$\overline{h}(x) + \epsilon^- \leq h(x) \leq \overline{h}(x) + \epsilon^+, \qquad x \in [0, L]$$

where $\overline{h}$ is a given function and ϵ^+, ϵ^- are plus and minus limits on the design parameter h. In some circumstances, these constraints do not allow the guide to be to much oversized.

on the mode exchange energy:

$$|A_j(x)|^2 \geq E_j^-(x) \text{ or } |A_j(x)|^2 \leq E_j^+(x), \qquad x \in [0, L]$$

where E_j^-, E_j^+ are minus and plus energy limits for the mode j. Such a constraint allows a limit on the generation or the attenuation of the mode j.

One can treat these quality constraints by an exterior penalty method of the cost function. For instance, to take the first quality constraints ($h \geq \overline{h} + \epsilon^-$) into account, the following penalized cost function is introduced:

$$J_\alpha(h) = J(h) + \alpha \int_0^L [h(x) - \overline{h}(x) - \epsilon]_-^2 \, dx, \qquad \text{with } \alpha > 0$$

where the function $[u]_-^2$ is defined by

$$[u]_-^2 = \begin{cases} u^2, & \text{if } u \leq 0 \\ 0, & \text{if } u \geq 0. \end{cases}$$

For the constraint to work well, the α coefficient must be chosen carefully.

In the following, in order to simplify this paper, we will omit these quality constraints. The dependence of the coupling matrices on the design parameter is generally strongly nonlinear. This means that the maximization problem (5) is not convex. Thus, no result of existence and uniqueness of (5) can be derived. The only result we can state is related to the existence of local maxima characterized by the Euler condition

$$\nabla_h J(h) = 0.$$

B. Approximated Optimum Design Problem

The design parameter h and (3) have to be discretized. We use a cubic spline approximation of the design parameter and a modified Crank–Nicolson scheme to discretize the equation.

Let us consider the following cubic spline approximation on a grid of points $(x_k)_{k=0,K}$ such as $x_0 = 0$ and $x_K = L$:

$$h_s(x) = \sum_{k=1}^{K} q_k(x) \chi_{[x_{k-1}, x_k]}(x) \tag{7}$$

where $q_k(x)$ is a cubic polynomial function and χ_I is the indicatrix function in the interval I (i.e., the function defined by $\chi_I(x) = 1$ if $x \in I$; $= 0$ otherwise).

From spline theory [13], there is an unique spline function $h_s \in C^2([0, L])$ satisfying the boundary conditions

$$h_s(x_k) = h_k \quad k = 0, K \quad \text{and} \quad h_s'(x_0) = h_s'(x_k) = 0.$$

Therefore, to take the constraints given by (6) into account, it is sufficient to choose as the unknown parameter the real vector

$$\mathbf{H} = (h_1, h_2, \cdots, h_{K-1})$$

associated to the spline function given by (7). This choice allows elimination of (6). It is of interest to note that the function $\mathbf{H} \rightarrow h_s$ is linear.

The standard Crank–Nicolson scheme [14] is not conservative. Thus, we use the following modified scheme on a regular grid $(x_n)_{n=0,\,N}$ with $x_n = n\Delta x$:

$$\begin{cases} \frac{\mathbf{A}^{n+1}-\mathbf{A}^N}{\Delta x} = \frac{1}{2}(\mathbf{C}^{n+1}+\mathbf{C}^n)(\mathbf{A}^{n+1}+\mathbf{A}^n), \\ \qquad\qquad\qquad\qquad\qquad\qquad n = 0,\, N-1 \\ \mathbf{A}^0 = \mathbf{A}_0 \end{cases} \quad (8)$$

with $\mathbf{C}^n = \mathbf{C}(h_s(x_n),\, x_n)$.

The scheme (8) is conservative, implicit, unconditionally stable and of second order. Because of the oscillating solution of (3), the conservation of the scheme appears to be essential. With these discretizations, the approximated optimum design problem has the following form:

$$\max \mathbf{H} \in \mathbf{R}^{K-1} J(\mathbf{H}), \qquad \text{with } J(\mathbf{H}) = |A_m^N|^2 \quad (9)$$

and becomes a standard finite dimension maximization problem.

C. Maximization Algorithms

The finite dimensional maximization problem (9) does not have any constraints. Thus, classical algorithms of ascent like the steepest ascent or Polak–Ribiere method [15] may be used to find local maxima of the cost function J characterized by: $\nabla J(\mathbf{H}) = 0$. For instance, the steepest ascent algorithm is

$$\begin{cases} \mathbf{H}^0 \text{ given} \\ \mathbf{H}^{p+1} = \mathbf{H}^p + \rho^p \nabla J(\mathbf{H}^P), \qquad p \geq 0 \end{cases} \quad (10)$$

where $\rho^p > 0$ is a step of ascent which is estimated by classical rules like dichotomy or the Goldstein method [15].

Remark: More efficient algorithms like quasi-Newton methods (e.g., BFGS methods [15]) are not recommended because of the strong nonconvexity of the problem.

The key point of optimization computation is the evaluation of the gradient. To make it as fast as possible, we use the well-known method based on the adjoint state. Let us explain it in the continuous case, which is less intricate than the discrete one. First of all, we note $W_k = \partial_{h_k}\mathbf{A}$ the derivative of $\mathbf{A}$ versus the kth design parameter. It is obvious that $\mathbf{W}_k$ is a solution of the following differential system:

$$\begin{cases} \frac{d}{dx}\mathbf{W}_k = C(\mathbf{H})\mathbf{W}_k + \partial_{h_k} C(\mathbf{H})\mathbf{A}, \qquad \forall x \in]0,\, L[\\ \mathbf{W}_k(0) = 0. \end{cases}$$

Moreover, we have

$$\partial_{h_k} J(\mathbf{H}) = 2\,\mathrm{Real}\,(\partial_{h_k} A_m(L),\, A_m(L))$$

where $(u,\, v)$ represents the Hermitian product between u and v.

Let us introduce the adjoint state vector $\mathbf{Z}$, solution of the following backward differential system:

$$\begin{cases} \frac{d}{dx}\mathbf{Z} = -C^*(\mathbf{H})\mathbf{Z}, \qquad \forall x \in]0,\, L[\\ \mathbf{Z}(L) = 2\mathbf{R}_m \end{cases}$$

with $\mathbf{R}_m = (0, \cdots, 0,\, A_m(L),\, 0, \cdots, 0)^t$.

Taking into account the two previous differential systems, the equality

$$\int_0^L \left(\frac{d}{dx}\mathbf{W}_k,\, \mathbf{Z}\right) dx = -\int_0^L \left(\mathbf{W}_k,\, \frac{d}{dx}\mathbf{Z}\right) dx + (\mathbf{W}_k(L),\, \mathbf{Z}(L))$$

leads to

$$\partial_{h_k} J(\mathbf{H}) = \mathrm{Real}\left[\int_0^L (\partial_{h_k}\mathbf{C}(\mathbf{H})\mathbf{A}(x),\, \mathbf{Z}(x))\, dx\right].$$

It is of interest to note that the gradient evaluation requires only one resolution of a backward differential system. This must be compared with the $(K-1)$ resolutions of the forward differential system in $\mathbf{W}_k$.

In the discrete case, similar, but more intricate calculations lead to

$$\partial_{h_k} J(\mathbf{H}) = \mathrm{Real}\left(\sum_{n=1}^{N} (\mathbf{A}^n,\, \partial_{h_k}(\mathbf{C}^n + \mathbf{C}^{n-1})^* \mathbf{Z}^n + \partial_{h_k}(\mathbf{C}^{n+1} + \mathbf{C}^n)^* \mathbf{Z}^{n+1})\right)$$

with $\mathbf{Z}^{N+1} = 0$, $\mathbf{Z}^1 = 0$ and $\mathbf{Z}^n$, $n = 2, N$ given by the adjoint scheme

$$\begin{cases} \frac{\mathbf{Z}^n - \mathbf{Z}^{n+1}}{\Delta x} = \frac{1}{4}[(\mathbf{C}^{n+1}+\mathbf{C}^n)^*\mathbf{Z}^{n+1} \\ \qquad\qquad +(\mathbf{C}^{n-1}+\mathbf{C}^n)^*\mathbf{Z}^n], \qquad n = 2,\, N-1 \\ \mathbf{Z}^N = \frac{1}{4}(\mathbf{C}^N + \mathbf{C}^{N+1})\mathbf{Z}^N + 2\mathbf{R}_m \end{cases}$$

where $\mathbf{R}_m = (0, \cdots, A_m^N, \cdots, 0)^t$. It appears that the discrete adjoint equation is not an obvious discretization of the continuous one. The derivatives $\partial_{h_k}\mathbf{C}^n$ are easy to obtain because the spline operator $\mathbf{H} \rightarrow h_s(x_n)$ is linear and the dependence of $\mathbf{C}^n$ on h_s is explicit.

D. Some Remarks on the Method

The strong nonconvexity of the cost function J is illustrated in Fig. 1 showing its variations as a function of the step ρ in a given direction

$$f(\rho) = J(\mathbf{H} - \rho\nabla J(\mathbf{H})).$$

One can see many maxima. Thus, there is no guarantee that the ascent method (10) gives a good solution to the problem (9), that is to say, a solution with the best possible efficiency.

Nevertheless, numerical experiments show that in most cases, the solution is very good since the initial parameter $\mathbf{H}^0$ does not privilege any situation. Moreover, as the function J oscillates less and less since the dimension of the design parameter decreases, a continuation method based on the

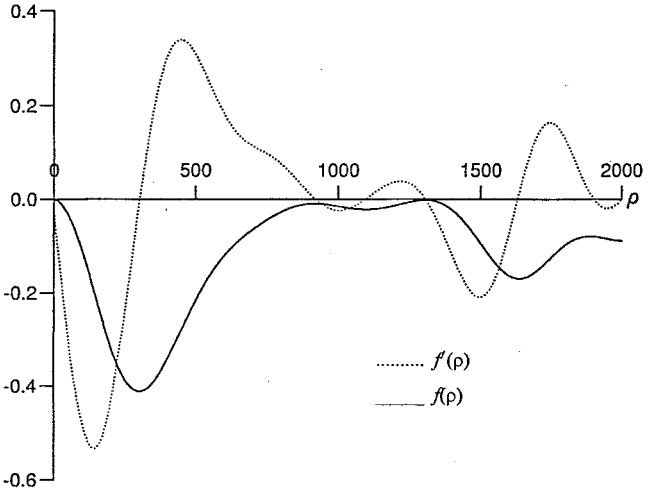

Fig. 1. Nonconvexity of the cost function J.

TABLE I

	Frequency (GHz)	Mode	Diameter (mm)	Length (mm)	Mode purity (%)
Taper-converter	100	TE_{01} to TE_{02}	38.1 to 27.8	550	99,8
Taper	110	TE_{64}	63.5 to 20	350	99
Converter	110	TE_{01} to TE_{11}	20	900	99,2

increase of the spline number of points allows a very good optimal solution in any realistic case. In order to speed up the numerical method, it is also possible to increase the number of modes (up to the cutoff) during the iterative process.

In this section, we have described the main ideas which have contributed to make an effective computational software of optimized waveguides. This software allows us to take into account ohmic losses, which are modelized with complex propagation constants. Moreover, sensitivities of waveguides to the design parameter perturbations or frequency perturbations can be estimated.

IV. Experimental Results

Three components have been designed and tested to characterize the operability of this method, in the various cases corresponding to the different coupling coefficients sets (see Table I). In the case of radius variations (Unger coefficients), a TE_{02} to TE_{01} mode converter, tapering from 27.8- to 38.1-mm diameter was designed to work at 100 GHz with 99.8% of mode purity. In the case of asymmetrical modes, an ultrashort TE_{64} waveguide, tapering from 63.5- to 20-mm-diameter taper was designed for operation at 110 GHz with 99% of mode purity. In the case of Morgan coefficients, a TE_{01} to TE_{11} mode 20-mm-diameter converter was designed at 110 GHz for a mode purity of 99.2%.

A. 100-GHz Taper-Converter [16]

This component illustrates the facilities offered by the method. The only data to be considered are the input and output modes, diameters, and slopes, the frequency, the length, and the desired mode purity. The code then runs until a shape satisfying these conditions is found. Thus, this method requires only a single calculation for the design of the taper-converter [5].

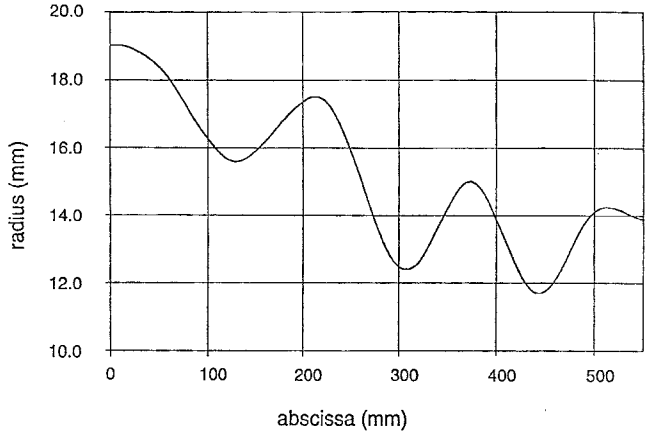

Fig. 2. Shape of the 100-GHz TE_{01} to TE_{02} taper-converter.

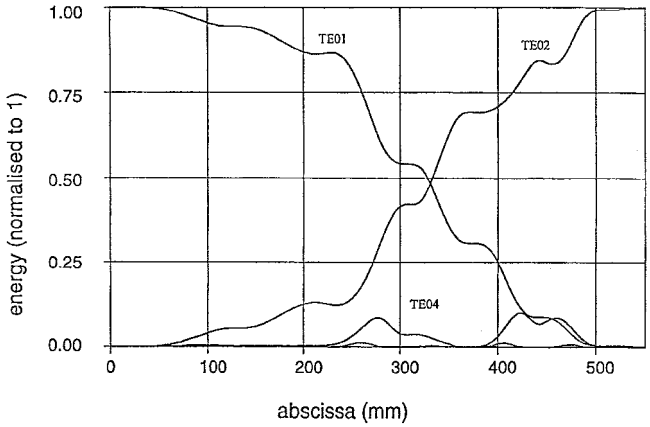

Fig. 3. Mode coupling for the 100-GHz TE_{01} to TE_{02} taper-converter.

The shape (Fig. 2) and the energy transfer (Fig. 3) curves are both regular, which is the sign of a noncritical case. Another sign is the weak coupling to higher modes. Should the component be much shorter, some singularities or abrupt variations would appear that would no longer satisfy the physical model approximations, and more energy would be transferred to higher modes.

B. 110-GHz Asymmetric Mode Taper

This component propagates the TE_{64} mode from 20-mm to 63.5-mm diameter at 110 GHz [17]. The final diameter is more than 20 times the free-space wavelength, and the final design is outstandingly short—350 mm. As the optimization method does not take into account evanescent modes, the considered modes are practically limited to those propagating in the smaller diameter of the component. Therefore, one has to select, among the various possible solutions, the smoother ones which are less likely to generate higher order modes. This was done by using penalty methods with constraints on the

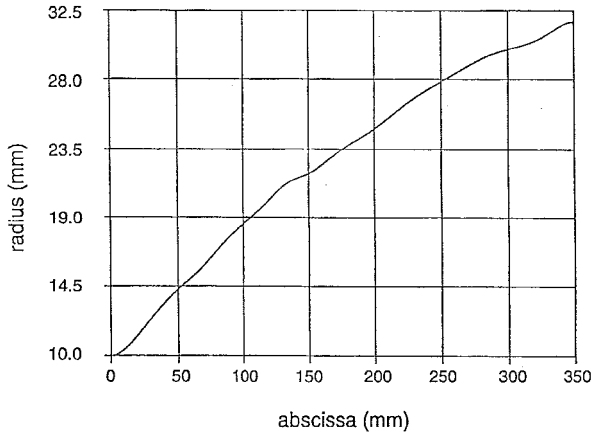

Fig. 4. Shape of the 100-GHz TE_{64} taper.

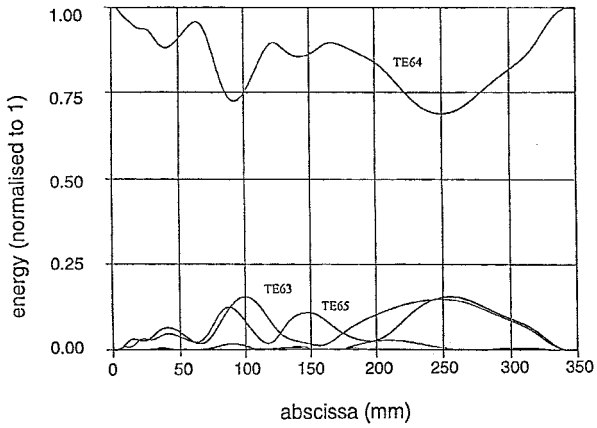

Fig. 5. Mode coupling for the 110-GHz TE_{64} taper.

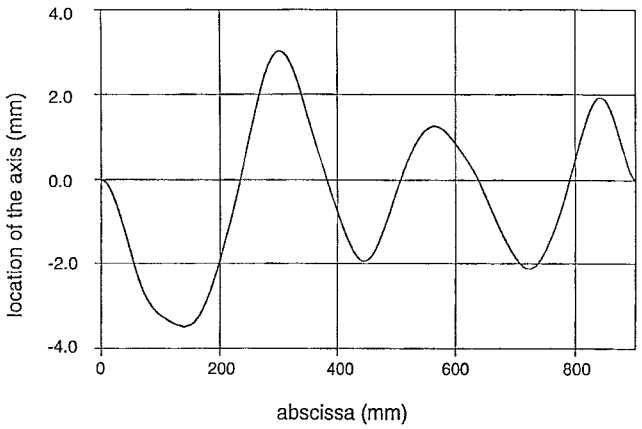

Fig. 6. Shape of the 110-GHz TE_{01} to TE_{11} converter.

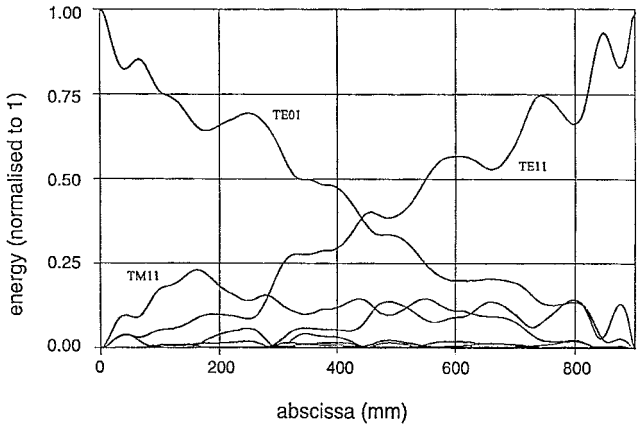

Fig. 7. Mode coupling for the 110-GHz TE_{01} to TE_{11} converter.

shape. The program was forced to choose a shape held within two given curves, and on the mode coupling. The program was forced to transfer no more than 30% of the energy on spurious modes.

A final calculation on the optimized structure, introducing a larger number of modes, enables one to check that the efficiency does not decrease when the number of modes increases. The chosen solution is shown in Fig. 4 for the guide radius and in Fig. 5 for the energy transfer curves.

C. 110 GHz TE_{01} *to* TE_{11} *Converter*

This "serpentine" converter has a constant diameter and its curvature varies with the abscissa in one plane. This is the condition for the generation of modes with a difference of azimuthal index of unity. If we consider the Fourier expansion of the deformation of the guide, the first harmonic corresponds to the beat-wavelength between the input and the output modes. In this case (Fig. 6), we have three periods, what seems to be the lower limit leading to acceptable smooth curvatures satisfying the physical-model approximations. The 99.2% computed mode purity (Fig. 7) also seems to be the upper limit in that case. Eleven modes were taken into account: TE_{01}, TE_{11}, TE_{12}, TE_{21}, TE_{22}, TE_{31}, TM_{11}, TM_{12}, TM_{21}, TM_{22}, and TM_{31}. The coupling to higher order modes proved to be quite negligible.

D. Experimental Setup

These components were measured with the following well-known over-moded waveguides measurement techniques: open-ended waveguide radiation patterns and k-spectrometer [18] patterns in both perpendicular and parallel polarities.

The microwave source was a Thomson TH42210 carcinotron delivering 1 W at 110 and 100 GHz. The detection was homodyne in the case of radiation measurements, and heterodyne with a 10-GHz intermediate frequency signal amplified by 30 dB in the case of the k-spectrometer measurements. Several intermediate tapers were designed with the same optimization techniques in order to provide good input mode purity at the input of the component under test.

Particular care was taken for the manufacturing of these components, leading to $+/-0.01$-mm maximum tolerances, either in the case of electroformed parts or by direct machining, especially in the case of the serpentine converter where the sometimes large curvatures required a careful programming of the milling machine.

E. Experimental Results

The measured patterns fit the theoretical predictions. Figs. 8–11, respectively, show the spectrometer and radiation patterns for the TE_{01} (at the input) and the TE_{02} (at the

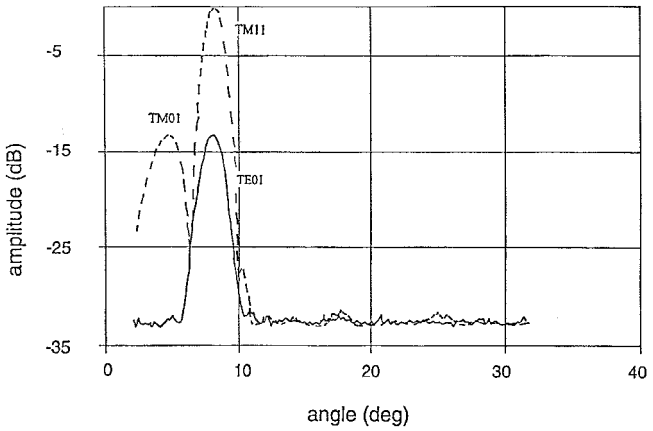

Fig. 8. K-spectrometer pattern for TE_{01} mode at the input of the 100-GHz taper-converter.

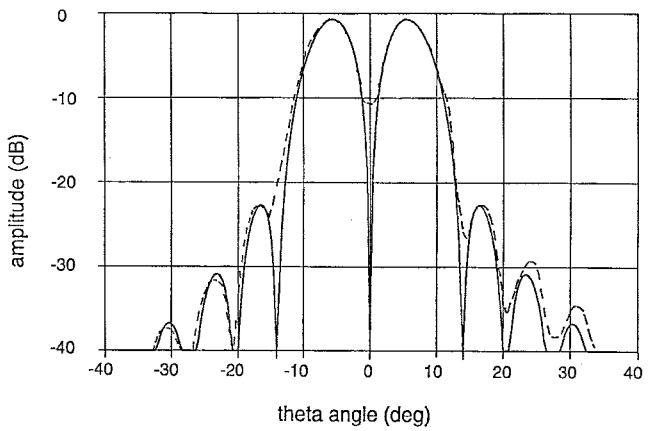

Fig. 9. Radiation pattern for TE_{01} mode at the input of the 100-GHz taper-converter.

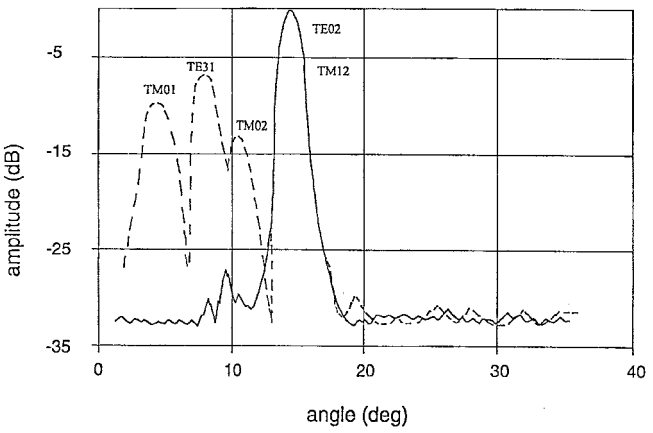

Fig. 10. K-spectrometer pattern for TE_{02} mode at the output of the 100-GHz taper-converter.

output) modes in the case of the 100-GHz taper-converter. Taking into account the k-spectrometer coupling coefficients, the spurious modes content is (TM_{01} 0.03%, TM_{11} 0.3%) at the input and $\mathrm{TE}_{01} < 0.2\%$, TE_{31} 0.01%, TM_{12} 0.05%, $\mathrm{TM}_{01} < 0.01\%$, $\mathrm{TM}_{02} < 0.01\%$) at the output. Figs. 12 and 13 show the radiation patterns of the TE_{11}, mode at the output of the serpentine converter.

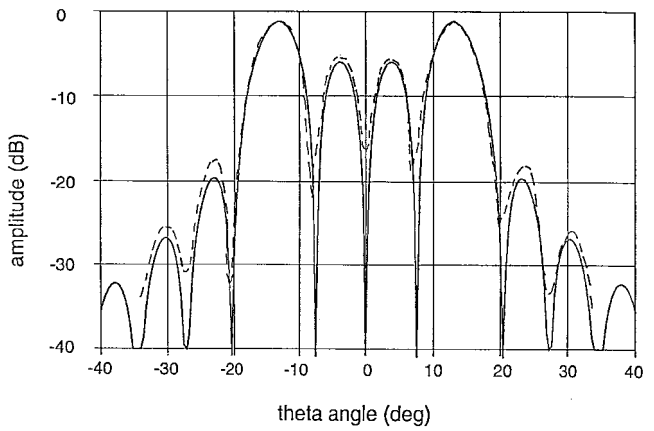

Fig. 11. Radiation pattern for TE_{02} mode at the output of the 100-GHz taper-converter.

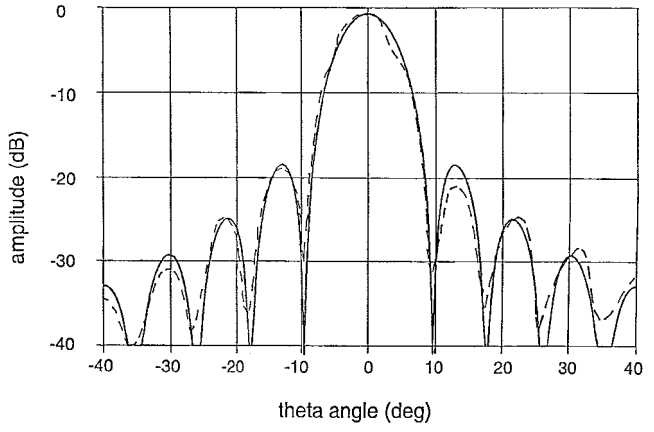

Fig. 12. E-plane radiation pattern for the output of the 110-GHz TE_{01} to TE_{11} converter.

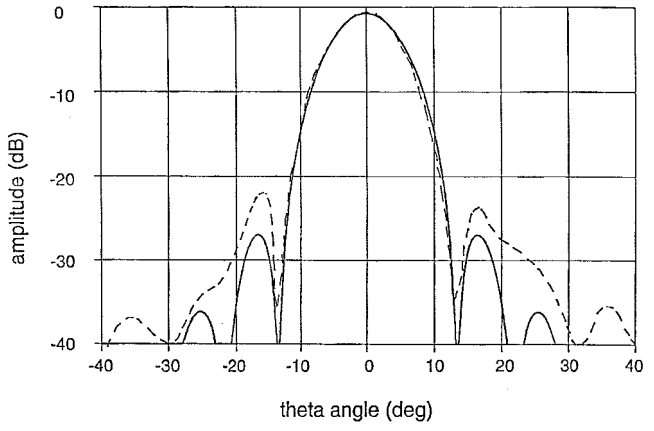

Fig. 13. H-plane radiation pattern for the output of the 110-GHz TE_{01} to TE_{11} converter.

V. Conclusion

An original method based on numerical optimization techniques has been developed which enables the calculation of ultrashort highly efficient multimode components. This method has been implemented and tested for both circular and asymmetrical modes in the case of longitudinal deformations of radius or curvature. The results show improvement of

components performances either for conversion efficiency or for total length. The ohmic losses can be taken into account and the sensitivity to mechanical errors through pseudorandom variations of dimensions can be evaluated. The behavior of the code is totally autonomous, and the result is generally a single final optimized structure in the case of moderately oversized components (transverse dimensions less than ten times the free-space wavelength). In the case of highly oversized components (transverse dimensions more than 20 times the free-space wavelength), there are generally several possible solutions, many of them being divergent with regard to the number of considered modes. For that reason, quality constraints methods have been developed, which allow elimination of these erratic solutions by limiting the slopes, the energy transfer and some other parameters. The next step of that work will concern helical converters and lateral outputs for high power devices. It will consist in bidimensional optimization applied to the Mourier coupling coefficients set.

REFERENCES

[1] J. L. Doane, "Polarization converters for circular waveguides modes," *Int. Electron.*, vol. 61, no. 6, pp. 1109–1133, 1986.

[2] M. Thumm, V. Erckmann, W. Kasparek, H. Kumric, G. A. Muller, P. G. Schuller, and R. Wilhelm, "Very high power mm-wave components in oversized waveguides," *Microwave J.*, pp. 103–121, Nov. 1986.

[3] H. Kumric, M. Thumm, and R. Wilhelm, "Optimization of mode converters for generating the fundamental TE_{01} mode from TE_{06} gyrotron output at 140 GHz," *Int. J. Electron.*, vol. 64, no. 1, pp. 77–94, 1988.

[4] H. Flügel and E. Kühn, "Computer-aided analysis and design of circular waveguide tapers," *IEEE Trans. Microwave Theory Tech.*, vol. 26, pp. 332–336, Feb. 1988.

[5] M. J. Buckley and R. J. Vernon, "Compact quasiperiodic and aperiodic TE_{0n} mode converters in overmoded circular waveguides for use with gyrotrons," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 712–721, June 1990.

[6] W. G. Lawson, "Theoretical evaluation of nonlinear tapers for a high-power gyrotron," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 1617–1622, Nov. 1990.

[7] J. L. Doane, "Propagation and mode coupling in corrugated and smooth-wall circular waveguides," *Infrared Millimeter Waves*, vol. 13, pp. 123–170, 1985.

[8] H. G. Unger, "Circular waveguide taper of improved design," *Bell Syst. Tech. J.*, pp. 899–912, July 1958.

[9] F. Sporleder and H. G. Unger, *Waveguides Tapers Transitions and Couplers.* Stenvenage, U.K.: Peregrinus, 1979.

[10] V. G. Pavel'ev, S. H. E. Tsimring, and V. E. Zapevalov, "Coupled cavities with mode conversion in gyrotrons," *Int. J. Electron.*, vol. 63, no. 3, pp. 379–391, 1987.

[11] S. P. Morgan, "Theory of curved circular waveguide containing an inhomogeneous dielectric," *Bell Syst. Tech. J.*, pp. 1209–1251, Sept. 1957.

[12] G. Mourier, G. Jendrzejczak, and E. Jedar, "Convertisseurs de modes en guides d'ondes circulaires surdimensionnés à 100 GHz," *J. Nationales Microondes*, June 22–24, 1984.

[13] G. Dahlquist and A. Björek, *Numerical Methods.* New York: Prentice-Hall, 1974.

[14] P. Henrici, *Discrete Variable Methods in Ordinary Differential Equations.* New York: Wiley, 1962.

[15] D. G. Luenberger, *Linear and Non-Linear Programming.* Reading, MA: Addison-Wesley, 1984.

[16] P. Garin, E. Giguet, J. M. Krieg, E. Lunéville, and G. Mourier, "Ultra compact taper mode converter design," presented at the *Proc. 15th Int. Conf. Infrared Millimeter Waves*, Orlando, FL, 1990.

[17] J. M. Krieg, E. Giguet, A. Dubrovin, P. Garin, and G. Mourier, "First results on a 110 GHz evacuated transmission line," presented at *Proc. 17th Int. Conf. Infrared Millimeter Waves*, Pasadena, CA, 1992.

[18] W. Kasparek and G. A. Müller, "The wavenumber spectrometer—An alternative to the directional coupler for multimode analysis in oversized waveguides," *Int. J. Electron.*, vol. 64, no. 1, pp. 5–20, 1988.

Eric Lunéville received the Ph.D. degree in applied mathematics from the Universite Pierre et Marie Curie, France, in 1988.

From 1988 to 1990, he held a Post-Doctoral position at Ecole des Mines de Paris, Paris, France. Since 1990, he has been the Maître de Conférences at the Universite Paris-Nord, Paris, France, where he has been working on optimization problems in wave propagation.

Jean-Michel Krieg received the engineering diploma from ENSHH, Besançon, France, in 1983, and the Ph.D. degree from the Université de Besançon, France, in 1988.

From 1987 to 1995, he was with Thomson Tubes Electroniques, Vélizy-Villacoublay, France, where he worked on high-power multimode transmission lines and the gyrotron. Since 1996, he has been a Research Engineer at the Observatoire de Paris, France, working on millimeter and submillimeter heterodyne receivers for radioastronomy.

Eric Giguet received the engineer diploma in electronics from ENSERG/INPG, Grenoble, France, in 1990.

From 1990 to 1991, he was with Thomson Tubes Electroniques, Vélizy-Villacoublay, France, where he worked on high-power multimode transmission lines and gyrotrons. From 1991 to 1993, he was with the Massachusetts Institute of Technology, Cambridge. In 1993, he returned to Thomson Tubes Electroniques, where he is presently a Gyrotron Projects Manager.