J. Mech. Phys. Solids Vol. 34, No. 4, pp. 333–358, 1986.
Printed in Great Britain.

0022–5096/86 $3.00+0.00
Pergamon Journals Ltd.

BOUNDS ON THE COMPLEX PERMITTIVITY OF A MULTICOMPONENT MATERIAL

K. GOLDEN†

Courant Institute, New York University, New York, NY 10012, U.S.A.

(*Received* 20 *February* 1985; *in revised form* 11 *November* 1985)

ABSTRACT

RECENTLY D. Bergman introduced a method for obtaining bounds on the effective dielectric constant (or conductivity) of a two-component medium. This method does not rely on a variational principle but instead exploits the properties of the effective parameter as an analytic function of the ratio of the component parameters. Here the method is extended to multicomponent media using techniques of several complex variables. We propose for the first time a series of bounds on the complex dielectric constant of a material of three or more components, as well as rederive the Wiener and Hashin-Shtrikman bounds for real parameters. In addition, we obtain in a simple manner a known infinite sequence of bounds for two-component media.

1. INTRODUCTION

DUE TO THE difficulty of calculating the effective parameters (e.g. dielectric constant, magnetic permeability, or electrical or thermal conductivity) of a heterogeneous material, there has been much interest in obtaining bounds on these parameters. WIENER (1912) gave optimal bounds on the effective parameters of a multicomponent material with fixed volume fractions and real component parameters. These bounds are sometimes known as the arithmetic and harmonic mean bounds. For isotropic materials, HASHIN and SHTRIKMAN (1962) improved Wiener's bounds using variational principles. Recently BERGMAN (1978, 1979, 1980a,b, 1981, 1982a,b, 1983) introduced a method for obtaining bounds on complex effective parameters which does not rely on variational principles. Instead it exploits the properties of the effective parameters as analytic functions of the component parameters. The method of Bergman has been elaborated upon in detail and applied to several problems by MILTON (1980, 1981a–c, 1982). A mathematical formulation of it was given by GOLDEN and PAPANICOLAOU (1983; hereafter referred to as GP1). However, aside from Bergman's trajectory approach which is discussed below, the method has been restricted to two-component materials, where the effective parameters are functions of a single complex variable, the ratio of the two component parameters. This analytic continuation method is extended to multicomponent media in a direct way for the first time by GOLDEN and PAPANICOLAOU (1985; hereafter referred to as GP2). There the analog for several

† Present address: Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, U.S.A.

complex variables of the single variable integral representation for effective parameters given in GP1 is obtained.

The main purpose here is to conjecture new complex versions of the Wiener (sketched in GP2) and Hashin–Shtrikman bounds for multicomponent media using the extended representation formula. In particular, in the Hashin–Shtrikman case we introduce a new fractional linear transformation of the effective parameters which diagonalizes to second order the perturbation expansion of the effective parameters about a homogeneous medium. The representation formula is applied to this new function to obtain the real Hashin–Shtrikman bounds and conjectured complex extensions, while it is directly applied to the effective parameter in the Wiener case. In addition, for two-component media, we derive in a simple manner in the Appendix a known infinite sequence of bounds which incorporates more and more information about the material, by using iterated fractional linear transformations.

In GP1 the integral representation involves a complex kernel containing the component parameter information and a positive measure containing information about the geometry of the composite. For three-component materials, the effective parameters are analytic functions of two complex variables. One of these two variables can be fixed as a multiple of the other, so that the effective parameters are treated as analytic functions of a single complex variable. BERGMAN (1978, 1983) has applied the single variable analytic method in this case to obtain the real Hashin–Shtrikman bounds. However, this approach makes the above-mentioned measure depend upon the component parameters as well as the geometry of the composite. A direct extension of the analytic continuation method should rely on a representation of the effective parameters which indeed separates the component parameters from the geometry of the composite. This is accomplished in GP2 and here by treating the effective parameters explicitly as analytic functions of several complex variables.

One reason for the usefulness of the multicomponent representation formula is the following. As mentioned above, the effective parameters can be expanded about a homogeneous medium. The information in this perturbation expansion can then be used along with the representation formula to continue the effective parameters beyond nearly homogeneous media to their full domain of analyticity.

2. OUTLINE OF METHODS AND RESULTS

We assume that the medium under study is an N-component microscopically isotropic dielectric. The mathematical details of our formulation of the multicomponent problem are given in PAPANICOLAOU and VARADHAN (1982), GP1 and GP2. Let $\varepsilon(x,\omega)$ be a stationary stochastic process in $x \in R^d$ and $\omega \in \Omega$, where Ω is the set of all realizations of our random medium. We write

$$\varepsilon(x,\omega) = \varepsilon_1\chi_1(x,\omega)+\varepsilon_2\chi_2(x,\omega)+\ldots+\varepsilon_N\chi_N(x,\omega), \tag{2.1}$$

where ε_j, $1 \leqslant j \leqslant N$, is the complex dielectric constant of medium j, and the indicator function $\chi_j(x,\omega)$ equals one for all realizations $\omega \in \Omega$ which have medium j at x, and equals zero otherwise. Let $E^k(x,\omega)$ and $D^k(x,\omega)$ be the stationary random electric and displacement fields satisfying

$$D_i^k(x,\omega) = \varepsilon(x,\omega)E_i^k(x,\omega) \tag{2.2}$$

$$\nabla\cdot D^k(x,\omega) = 0 \tag{2.3}$$

$$\nabla\times E^k(x,\omega) = 0 \tag{2.4}$$

$$\int_\Omega P\,(\mathrm{d}\omega)\; E^k(x,\omega) = e_k, \tag{2.5}$$

where e_k is a unit vector in the kth direction for some $k = 1,2,\ldots,d$. In (2.5) P is a probability measure on Ω which is compatible with the stationarity of the problem (PAPANICOLAOU and VARADHAN, 1982), so that we may focus attention at say $x = 0$, and then drop the x notation.

The effective dielectric constant ε_{ik}^* may now be defined as

$$\varepsilon_{ik}^* = \int_\Omega P\,(\mathrm{d}\omega)\; D_i^k(\omega). \tag{2.6}$$

It is shown in GP1 that this ensemble average in an infinite stationary medium coincides with the more standard definition involving a volume average. Since (2.2)–(2.6) are linear in $\varepsilon(\omega)$, ε_{ik}^* depends only on the ratios $h_i = \varepsilon_i/\varepsilon_N$, $i = 1,\ldots,N-1$. We define

$$m_{ik}(h_1,\ldots,h_{N-1}) = \frac{\varepsilon_{ik}^*}{\varepsilon_N} = \int_\Omega P\,(\mathrm{d}\omega)\left(\sum_{j=1}^{N-1} h_j\chi_j(\omega)+\chi_N(\omega)\right)E_i^k(\omega). \tag{2.7}$$

Clearly m_{ik} has the same domain of analyticity in C^{N-1} as does E_i^k. It can be shown (GOLDEN, 1984) that if at a finite $(h_1,\ldots,h_{N-1})$ there exists a unique solution E_i^k to a suitable formulation of (2.3)–(2.5), then E_i^k is analytic at $(h_1,\ldots,h_{N-1})$. The condition for existence and uniqueness is that the smallest convex set containing $\{1, h_1,\ldots,h_{N-1}\}$ does not contain the origin in C (GP2). Thus m_{ik} is analytic when $(h_1,\ldots,h_{N-1})$ satisfies this condition. Furthermore, from the symmetric form of the definition

$$\varepsilon_{ik}^* = \int_\Omega P\,(\mathrm{d}\omega)\sum_{j=1}^{d}\varepsilon(\omega)\,E_j^k(\omega)\,\overline{E_j^i(\omega)}, \tag{2.8}$$

where the bar denotes complex conjugation, it is apparent that the diagonals m_{kk} map $\{\mathrm{Im}\,h_1 > 0\}\times\ldots\times\{\mathrm{Im}\,h_{N-1} > 0\}$ into the upper half plane with $m_{kk}(\bar{h}_1,\ldots,\bar{h}_{N-1}) = \overline{m_{kk}(h_1,\ldots,h_{N-1})}$.

We now outline the method which exploits these properties of $m = m_{kk}$, and briefly describe the resulting bounds. As mentioned in Sect. 1, the analytic continuation method relies on a representation of m which separates its dependence on the h_i's from the geometry of the mixture. This representation is obtained by treating m as an analytic function on a product of upper half planes. It is mathematically convenient to consider an auxiliary function $f(\zeta_1,\ldots,\zeta_{N-1}) = i(1-m(h_1,\ldots,h_{N-1}))$, $i = \sqrt{-1}$, on a product of discs $\{|\zeta_i| < 1\}$, where it has positive real part, and h_i and ζ_i are related by a standard conformal mapping. Then the above representation for f amounts to a manipulated form of the Cauchy Integral Formula applied to a product of discs, and is displayed in (4.1), (4.2) and (4.3). The parameter information (h_i) is contained in the kernel involving the H_i and the geometrical information is contained in the positive measure μ. This measure μ is concentrated on the (distinguished) boundary of the product of discs and arises from the boundary values of the real part of f. As such, μ cannot generally be identified with, say, a smooth function, but has singular ("δ-function") components. Constraints on μ coming from assumptions about the mixture geometry (e.g. volume fractions) are imposed by relating a suitable form of (4.1) to the perturbation expansion (5.3) mentioned in Sect. 1.

The extremization procedure which yields the bounds is based on the following two observations: (i) for fixed h_i, $1 \leqslant i \leqslant N-1$, (4.1) is linear in μ, and (ii) the class of admissible μ forms a compact, convex set M. Then for fixed h_i, extremal values of m (or f) are attained by extreme points of M, just as in linear programming. We have found that the multicomponent Wiener and Hashin–Shtrikman expressions for real parameters arise from the simplest extreme points of M, which are products of a "δ-function" in one direction with uniform measure in the other directions, as in (5.1) for three-component media. We then use the same type of simple measures to propose new complex bounds, which for three-component media are illustrated in Fig. 1. For the outermost bound (arcs a, b, c), only knowledge of ε_1, ε_2 and ε_3 is assumed. The volume fractions are assumed as well to obtain arcs d and e, while statistical isotropy is further assumed for arcs f and g. Arcs a, b, and c are attained by actual composite materials, as are parts of arcs d and e. We do not know if arcs f and g are optimal.

The fact that the above arcs are images of extreme points of admissible classes of μ does not provide mathematical proof that the obtained regions form rigorous bounds. This difficulty arises (for $N \geqslant 3$) because the full set of extreme points of M is much larger than the simple class that we use, and, in fact, has not been completely characterized (RUDIN, 1970, 1983; MCDONALD, 1982). Our extremization procedure (to second order in the perturbation expansion) is thus stated in the form of five function theoretic hypotheses which are as yet unproven in the above context. Nevertheless, using an extended version of Bergman's trajectory method, MILTON (1984) has proven subsequently to the present work that arcs a, b, and c and arcs d and e do indeed form bounds for the complex dielectric constant. That arcs f and g form a bound is still conjecture.

3. BOUNDS FOR TWO-COMPONENT MEDIA

Let $h = \varepsilon_1/\varepsilon_2$, $s = 1/(1-h)$, and $F_{ik}(s) = \delta_{ik} - m_{ik}(h)$. From section 2, $m_{ik}(h)$ is analytic off the negative real axis $(-\infty, 0]$, or $F_{ik}(s)$ is analytic off $[0, 1]$. In GP1, it was proved that there exist finite Borel measures μ_{ik} (dz) on $[0, 1)$ such that the diagonals μ_{kk} (dz) are positive and

$$F_{ik}(s) = \int_0^1 \frac{\mu_{ik}\,(\mathrm{d}z)}{s-z}, \quad i, k = 1, \ldots, d, \quad s \notin [0, 1]. \tag{3.1}$$

One proof of (3.1) depends on the operator representation arising from (2.3),

$$F_{ik}(s) = \int_\Omega P\,(\mathrm{d}\omega)\,\chi_1[(s+\Gamma\chi_1)^{-1}\,e_k]\cdot e_i, \tag{3.2}$$

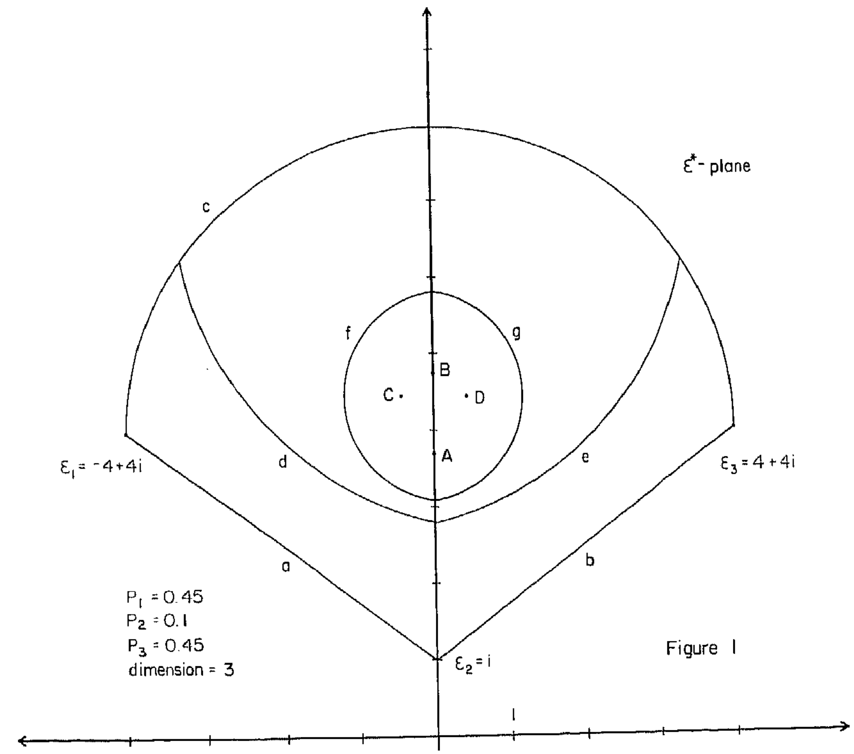

FIG. 1. Various points and arcs in the complex ε^*-plane. a, $L_{12}(\alpha)$ in (6.4); b, $L_{23}(\alpha)$ in (6.4); c, $C_{13}(\alpha)$ in (6.5); d, $B_3(z_1)$ in (6.12); e, $B_1(z_1)$ in (6.13); f, circular boundary of R^*_G in (6.23) and (6.26); g, circular boundary of $R^*_{\tilde{G}}$ similar to (6.23) and (6.26); A, $\varepsilon^* = p_1\varepsilon_1 + p_2\varepsilon_2 + p_3\varepsilon_3$; B, $\varepsilon^* = 1/(p_1/\varepsilon_1 + p_2/\varepsilon_2 + p_3/\varepsilon_3)$; C, $\varepsilon^* = \varepsilon_3 + 1/(1/A_3 - 1/3\varepsilon_3)$ as in (5.28); D, $\varepsilon^* = \varepsilon_1 + 1/(1/A_1 - 1/3\varepsilon_1)$ as in (5.28).

where $\Gamma = \nabla(-\Delta)^{-1}\nabla\cdot$ (and the differential operators $\partial/\partial x_i$ in Γ are replaced by the infinitesimal generators of the translation group on Ω). In the Hilbert space $L^2(\Omega, P)$ with weight $\chi_1(\omega)$ in the inner product, $\Gamma\chi_1$ is a bounded self adjoint operator of norm less than or equal to one (GP1). The formula (3.1) is the spectral representation of the resolvent $(s+\Gamma\chi_1)^{-1}$, where μ_{ik} (dz) is the spectral measure of the family of projections of $\Gamma\chi_1$. Another related proof exploits the fact that $-F_{kk}(s)$ has positive imaginary part when $\operatorname{Im} s > 0$, and is analytic at $s = \infty$. Then a general representation theorem in function theory (AKHIEZER and GLAZMAN, 1966) gives (3.1) for the diagonals $i = k$. It is this function theory approach that we use to extend the analytic continuation method to multicomponent media.

For $|s| > 1$, (3.1) can be expanded about a homogeneous medium ($s = \infty$ or $h = 1$),

$$F_{ik}(s) = \frac{\mu^{(0)}_{ik}}{s} + \frac{\mu^{(1)}_{ik}}{s^2} + \frac{\mu^{(2)}_{ik}}{s^3} + \ldots, \quad \mu^{(n)}_{ik} = \int_0^1 z^n \mu_{ik}\,(\mathrm{d}z). \tag{3.3}$$

Equating (3.3) to the same expansion of (3.2) yields

$$\mu^{(n)}_{ik} = (-1)^n \int_\Omega P\,(\mathrm{d}\omega)\,[\chi_1(\Gamma\chi_1)^n e_k]\cdot e_i. \tag{3.4}$$

When $i = k$ the moments $\mu^{(n)}_{kk}$ uniquely determine the positive μ_{kk}. Then (3.1) provides the analytic continuation of (3.3) to the full complex s-plane excluding $[0, 1]$. When $i \neq k$, μ_{ik} is a signed measure of mass 0.

We now focus on one diagonal coefficient ε^*_{kk} and call it ε^*, with $m = \varepsilon^*/\varepsilon_2$ and

$$F(s) = 1 - m(h) = \int_0^1 \frac{\mu\,(\mathrm{d}z)}{s-z}, \quad s \notin [0, 1]. \tag{3.5}$$

Bounds on ε^* are obtained as follows. By (3.4) the mass of μ in (3.5) equals the volume fraction p_1 of medium 1, which is less than or equal to 1. For $s \in C$ off [0, 1], $F(s, \mu)$ in (3.5) is a linear functional from the set M of positive measures of mass $\leqslant 1$ on [0, 1] into C. Thus extreme points of the set of values of $F(s, \mu)$ in C are attained by one-point measures $\alpha\delta_a\,(\mathrm{d}z)$, $0 \leqslant \alpha$, $a \leqslant 1$ since they are the extreme points of M (DUNFORD and SCHWARTZ, 1958). For these measures F has the form

$$F(s) = \frac{\alpha}{s-a}, \quad 0 \leqslant \alpha \leqslant 1, \quad 0 \leqslant a \leqslant 1. \tag{3.6}$$

The condition $F(1) \leqslant 1$ (BERGMAN, 1978) determines the allowed region in the F-plane. It is the image of the triangle in (α, a)-space, defined by $\alpha + z \leqslant 1$, $0 \leqslant \alpha \leqslant 1$, $0 \leqslant a \leqslant 1$, under the mapping (3.6). The region is bounded by a circular arc $C(\alpha)$ and a line segment $L(\alpha)$, with

$$C(\alpha) = \frac{\alpha}{s-(1-\alpha)}, \quad L(\alpha) = \frac{\alpha}{s}, \quad 0 \leqslant \alpha \leqslant 1. \tag{3.7}$$

These bounds are optimal and can be attained by slab composites aligned perpendicular and parallel to the applied field. The arcs are traced out as the volume fraction varies. For real parameters with $\varepsilon_1 \leqslant \varepsilon_2$ these bounds collapse to the interval $\varepsilon_1 \leqslant \varepsilon^* \leqslant \varepsilon_2$.

If the volume fractions p_1 and $p_2 = 1 - p_1$, are fixed as well as $s \in C$, then the mass of μ in (3.5) is fixed with $\mu^{(0)} = p_1$. Then the values of F lie inside the circle parameterized by

$$C_1(z) = \frac{p_1}{s-z}, \quad -\infty \leqslant z \leqslant \infty. \tag{3.8}$$

On the other hand,

$$E(s) = 1 - \varepsilon_1/\varepsilon^* = \frac{1 - sF(s)}{s(1-F(s))} \tag{3.9}$$

is an upper half plane function with the same domain of analyticity as $F(s)$ (BERGMAN, 1982b) so that it has a representation like (3.5),

$$E(s) = \int_0^1 \frac{\nu\,(\mathrm{d}z)}{s-z}, \quad s \notin [0, 1]. \tag{3.10}$$

The perturbation expansion of $E(s)$ forces the mass of ν equal to p_2. Then the values of $E(s)$ lie inside the circle

$$\hat{C}_1(z) = \frac{p_2}{s-z}, \quad -\infty \leqslant z \leqslant \infty. \tag{3.11}$$

In the ε^*-plane the intersection of these two regions is bounded by two circular arcs corresponding to $0 \leqslant z \leqslant p_2$ in (3.8) and $0 \leqslant z \leqslant p_1$ in (3.11). As BERGMAN (1982b, 1983) and MILTON (1981b) show, these bounds are optimal. They are attained by a composite of uniformly aligned spheroids of material 1 in all sizes coated with confocal shells of material 2, and vice versa. The arcs are traced out as the aspect ratio varies.

When ε_1 and ε_2 are real and positive the region collapses to the interval $1/(p_1/\varepsilon_1 + p_2/\varepsilon_2) \leqslant \varepsilon^* \leqslant p_1\varepsilon_1 + p_2\varepsilon_2$, which are the Wiener bounds. They are attained by parallel slabs of the materials. The upper Wiener bound is equivalent to a lower bound on F obtained by setting $z = 0$ in (3.8), and the lower Wiener bound is the same as a lower bound on E in (3.11) with $z = 0$ (or equivalently the upper bound on F in (3.8) with $z = p_2$, as required by $F(1) \leqslant 1$).

If the material is further assumed to be statistically isotropic, then $\mu^{(1)}$ in (3.4) can be

computed, with $\mu^{(1)} = p_1 p_2/d$, where d is dimension, so that F is known to second order

$$F(s) = \frac{p_1}{s} + \frac{p_1 p_2}{ds^2} + \ldots. \tag{3.12}$$

A convenient way of including this information is to use the transformation (BERGMAN, 1982b)

$$F_1(s) = \frac{1}{p_1} - \frac{1}{sF(s)}. \tag{3.13}$$

The function $F_1(s)$ is an upper half plane function analytic off [0, 1] so that it has the representation

$$F_1(s) = \int_0^1 \frac{\mu_1\,(\mathrm{d}z)}{s-z}. \tag{3.14}$$

Under (3.12) F_1 is known only to first order

$$F_1(s) = \frac{p_2/p_1 d}{s} + \ldots, \tag{3.15}$$

which forces $\mu_1^{(0)} = p_2/p_1 d$. Then the values of $F_1(s)$ lie inside the circle $p_2/(p_1 d(s-z))$, $-\infty \leqslant z \leqslant \infty$. Since F is fractional linear in F_1, this circle is transformed to a circle in the F-plane

$$C_2(z) = \frac{p_1(s-z)}{s(s-z-p_2/d)}, \quad -\infty \leqslant z \leqslant \infty. \tag{3.16}$$

Applying similar considerations to

$$E(s) = \frac{p_2}{s} + \frac{p_1 p_2(d-1)}{ds^2} + \ldots \tag{3.17}$$

gives a circle in the E-plane,

$$\hat{C}_2(z) = \frac{p_2(s-z)}{s(s-z-p_1(d-1)/d)}, \quad -\infty \leqslant z \leqslant \infty. \tag{3.18}$$

In the ε^*-plane the intersection of these two circular regions is bounded by two circular arcs corresponding to $0 \leqslant z \leqslant (d-1)/d$ in (3.16) and $0 \leqslant z \leqslant 1/d$ in (3.18). The vertices of the region, $C_2(0) = p_1/(s - p_2/d)$ and $\hat{C}_2 = p_2/(s - p_1(d-1)/d)$ are attained by the Hashin–Shtrikman geometries (spheres of all sizes of material 1 in the volume fraction p_1 coated with spherical shells of material 2 in the volume fraction p_2, and vice versa), and lie on the arcs of the first order bounds. While there are at least five points on the arc in (3.16) that are attainable (MILTON, 1981b), the arc in (3.18) violates the interchange inequality (KELLER, 1964; SCHULGASSER, 1976)

$$m(h)m(1/h) \geqslant 1, \tag{3.19}$$

which becomes an equality in two dimensions, and is consequently not optimal. MILTON (1980) and BERGMAN (1982b) have improved this bound by incorporating (3.19).

When ε_1 and ε_2 are real and positive the region collapses to the interval,

$$\varepsilon_1 + p_2 \Big/ \left(\frac{1}{\varepsilon_2 - \varepsilon_1} + \frac{p_1}{d\varepsilon_1}\right) \leqslant \varepsilon^* \leqslant \varepsilon_2 + p_1 \Big/ \left(\frac{1}{\varepsilon_1 - \varepsilon_2} + \frac{p_2}{d\varepsilon_2}\right), \quad \varepsilon_1 \leqslant \varepsilon_2, \tag{3.20}$$

which are the Hashin–Shtrikman bounds. The upper bound is equivalent to a lower bound on F_1 obtained by setting $\mu_1 = (p_2/p_1 d)\delta_0$ in (3.14). Similarly, the lower bound is equivalent to a lower bound on $E_1 = 1/p_2 - 1/sE$.

The higher moments $\mu^{(n)}$, $n \geqslant 2$ depend on $(n+1)$-point correlation functions and cannot be calculated in general, although the interchange inequality forces relations among them (MILTON, 1981a; GOLDEN, 1984). If $\mu^{(0)}, \ldots, \mu^{(n)}$ are known, then the transformation (3.13) can be iterated to produce a function of the same type as F, known only to first order, which is then easily extremized. This iteration procedure is described in the Appendix. The resulting bounds form a nested sequence of lens-shaped regions.

The transformation F_1 in (3.13) can be used to obtain the bounds of TARTAR and MURAT (1981) and LURIE and CHERKAEV (1984) for anisotropic composites. In two dimensions, following (3.3) we write

$$F_{11}(s) = \frac{p_1}{s} + \frac{\beta_1}{s^2} + \ldots, \quad F_{22}(s) = \frac{p_1}{s} + \frac{\beta_2}{s^2} + \ldots. \tag{3.21}$$

The anisotropic interchange inequality (MENDELSON, 1975; KOHLER and PAPANICOLAOU, 1982)

$$m_{11}(h)m_{22}(1/h) = 1 \tag{3.22}$$

implies that (GOLDEN, 1984)

$$\beta_1 + \beta_2 = p_1 p_2. \tag{3.23}$$

Equation (3.23) can also be obtained directly from (3.4), where $\beta_1 = \mu_{11}^{(1)}$and and $\beta_2 = \mu_{22}^{(1)}$. The new functions

$$\tilde{F}_{11} = \frac{1}{p_1} - \frac{1}{sF_{11}}, \quad \tilde{F}_{22} = \frac{1}{p_1} - \frac{1}{sF_{22}} \tag{3.24}$$

are of the same type as F_{11} and F_{22} and are known to first order

$$\tilde{F}_{11} = \frac{\beta_1/p_1^2}{s} + \ldots, \quad \tilde{F}_{22} = \frac{\beta_2/p_1^2}{s} + \ldots. \tag{3.25}$$

Applying the standard minimization for a fixed real s gives

$$\tilde{F}_{11}(s) + \tilde{F}_{22}(s) \geqslant \frac{p_2/p_1}{s}, \tag{3.26}$$

which is equivalent to

$$\frac{1}{F_{11}} + \frac{1}{F_{22}} \leqslant \frac{2s - p_2}{p_1}. \tag{3.27}$$

Similar considerations can be applied to E, giving

$$\frac{1}{E_{11}} + \frac{1}{E_{22}} \leqslant \frac{2s - p_1}{p_2}. \tag{3.28}$$

The bounds (3.27) and (3.28) are optimal. This anisotropic case was covered previously using a similar transformation by MILTON and GOLDEN (1985).

4. THE POLYDISC REPRESENTATION FORMULA

For simplicity we consider three-component media so that $m(h_1, h_2) = \varepsilon^*/\varepsilon_3$ and $F(s_1, s_2) = 1 - m(h_1, h_2)$ are functions of two complex variables, $h_1 = \varepsilon_1/\varepsilon_3$, and $h_2 = \varepsilon_2/\varepsilon_3$, or $s_1 = 1/(1-h_1)$ and $s_2 = 1/(1-h_2)$. Denote $U^2 = \{\mathrm{Im}\, s_1 > 0\} \times \{\mathrm{Im}\, s_2 > 0\}$. As a counterpart of $F(s_1, s_2)$: $U^2 \to \{\mathrm{Im}\, F < 0\}$, consider $f(\zeta_1, \zeta_2) = iF(s_1, s_2)$ with $f(\zeta_1, \zeta_2)$: $D^2 \to \{\mathrm{Re}\, f > 0\}$ where $D^2 = \{|\zeta_1| < 1\} \times \{|\zeta_2| < 1\}$ and $\zeta_j = (s_j - i)/(s_j + i)$, $j = 1, 2$. In GP2 the analogue of (3.5) for f in D^2 is derived. The result may be stated as follows. A function $f(\zeta_1, \zeta_2)$ which is holomorphic with positive

real part in D^2 may be represented as

$$f(\zeta_1, \zeta_2) = iv(0,0) + \frac{1}{2}\int_0^{2\pi}\int_0^{2\pi} (H_1 H_2 + H_1 + H_2 - 1)\,\mu(\mathrm{d}t_1, \mathrm{d}t_2), \tag{4.1}$$

where v is the imaginary part of f,

$$H_1 = \frac{e^{it_1} + \zeta_1}{e^{it_1} - \zeta_1}, \quad H_2 = \frac{e^{it_2} + \zeta_2}{e^{it_2} - \zeta_2} \tag{4.2}$$

and μ is a positive Borel measure satisfying

$$\int_0^{2\pi}\int_0^{2\pi} e^{i(nt_1 + mt_2)}\mu(\mathrm{d}t_1, \mathrm{d}t_2) = 0 \quad \text{when} \quad nm < 0, \quad n, m \in \{0, \pm 1, \pm 2, \ldots\}. \tag{4.3}$$

The representation (4.1) is obtained by manipulation of the two-variable Cauchy formula for $f(\zeta_1, \zeta_2)$ in D^2. The measure $\mu(\mathrm{d}t_1, \mathrm{d}t_2)$ in (4.1) and (4.3) is the radial (weak*) limit of $u(Re^{it_1}, Re^{it_2})\,\mathrm{d}t_1\mathrm{d}t_2$ as $R \to 1$, where u is the real part of f. The Fourier condition (4.3) arises from the fact that f has a Fourier series with only nonnegative powers of $Re^{i\theta_1}$ and $Re^{i\theta_2}$.

Because $F(s_1, s_2)$ is analytic when s_1 and s_2 are real and off $[0, 1]$, the measure $\mu(\mathrm{d}t_1, \mathrm{d}t_2)$ must vanish on a corresponding subset E of $T^2 = \{0 \leqslant t_1 < 2\pi\} \times \{0 \leqslant t_2 < 2\pi\}$. The effect on μ of this support condition and the Fourier condition (4.3) is discussed in the author's thesis (GOLDEN, 1984).

5. THE WIENER AND HASHIN–SHTRIKMAN BOUNDS FOR THREE-COMPONENT MEDIA

For two-component media the bounds were obtained by examining the images of extreme points of the set of positive measures of mass $\leqslant 1$ under the mapping (3.5). Denote M_1 = {positive Borel measures μ on T^2 that satisfy the Fourier condition (4.3) and have total mass $\leqslant 1$}. The simplest extreme points of M_1 have the form

$$\mu_1^* = \alpha\delta_{t_1^*}(\mathrm{d}t_1) \times \frac{\mathrm{d}t_2}{2\pi}, \quad \mu_2^* = \beta\delta_{t_2^*}(\mathrm{d}t_2) \times \frac{\mathrm{d}t_1}{2\pi}, \tag{5.1}$$

where $0 \leqslant t_1^*$, $t_2^* < 2\pi$ and $\alpha, \beta \leqslant 1$. However, as mentioned in Sect. 2, the full set of extreme points of M_1 has not been completely characterized. Nevertheless, we have been able to recover the Wiener and Hashin–Shtrikman bounds for three-component media with real ε_1, ε_2, and ε_3 by using sums of the measures in (5.1) with appropriate weights. It appears then that such measures give extremal values of functions represented by (4.1) for fixed ζ_1 and ζ_2. We will make this conjecture more precise in the present and following sections.

Before we begin with the derivation of the bounds, we should point out the following feature of our procedure. The representation (4.1) provides a map from the set of random geometries of a three-component material to M_1. However, not every measure in M_1, even those with the appropriate mass restriction corresponds to a random geometry. This is so because each of the six related functions $1 - \varepsilon^*/\varepsilon_3$, $1 - \varepsilon_3/\varepsilon^*$,

$1-\varepsilon^*/\varepsilon_2$, $1-\varepsilon_2/\varepsilon^*$, $1-\varepsilon^*/\varepsilon_1$, and $1-\varepsilon_1/\varepsilon^*$ for an appropriate set of variables, has a counterpart in D^2 with the representation (4.1) and the same mass restriction on μ. Then, given any one function, say $F(s_1, s_2) = 1-\varepsilon^*/\varepsilon_3$, the measure μ in M_1 that represents it must contain this information, thereby restricting the admissible class. The practical significance of this fact is that the full boundary of the allowed region for ε^* under various assumptions cannot be obtained by using the measures in (5.1) in the representation formula for $F = 1-\varepsilon^*/\varepsilon_3$ alone. We must consider one or more of the above auxiliary functions as well as F. In other words, a bound obtained from consideration of $1-\varepsilon_3/\varepsilon^*$ using the measures in (5.1) corresponds to a bound on $1-\varepsilon^*/\varepsilon_3$ involving measures much more complicated than those in (5.1). Therefore, we employ the above auxiliary functions as an integral part of our procedure.

In order to rederive the Wiener and Hashin–Shtrikman bounds, we give the analogue of (3.2) with $i = k$ for three-component media,

$$F(s_1, s_2) = \int_\Omega P(\mathrm{d}\omega)\left(\frac{1}{s_1}\chi_1 + \frac{1}{s_2}\chi_2\right)\left[\left(I + \frac{1}{s_1}\Gamma\chi_1 + \frac{1}{s_2}\Gamma\chi_2\right)^{-1} e_k\right]\cdot e_k, \tag{5.2}$$

where I is the identity. Note that the operators $\Gamma\chi_1$ and $\Gamma\chi_2$ in (5.2) do not commute, so that it is not immediately clear how to extend the spectral analysis described at the beginning of Sect. 3 to multicomponent media. For $|s_1| > 1$ and $|s_2| > 1$, (5.2) can be expanded about a homogeneous medium ($s_1 = s_2 = \infty$),

$$F(s_1, s_2) = \int_\Omega P(\mathrm{d}\omega)\left[\left(\frac{\chi_1}{s_1} + \frac{\chi_2}{s_2} - \frac{\chi_1\Gamma\chi_1}{s_1^2} - \frac{\chi_2\Gamma\chi_2}{s_2^2} - \frac{(\chi_1\Gamma\chi_2 + \chi_2\Gamma\chi_1)}{s_1 s_2} + \cdots\right)e_k\right]\cdot e_k. \tag{5.3}$$

To state the first form of our conjecture precisely, we introduce a function $K(s_1, s_2)$ that has the same properties as $F(s_1, s_2)$. Namely,

$$\begin{aligned}
&\text{(i) } K \text{ is analytic in } U^2\\
&\text{(ii) } K\colon U^2 \to \{\operatorname{Im} K < 0\}\\
&\text{(iii) } K(\bar{s}_1, \bar{s}_2) = \bar{K}(s_1, s_2)\\
&\text{(iv) } K(\infty, \infty) = 0\\
&\text{(v) } K(s_1, s_2) \text{ is analytic for real } s_1 \text{ and } s_2 \text{ when both } s_1 \text{ and } s_2 \text{ are off } [0, 1].
\end{aligned} \tag{5.4}$$

Because K has these properties, it has a counterpart in D^2 which has the representation (4.1). With sums of the measures in (5.1) in mind, for the counterpart of K in D^2 we let $\mu(\mathrm{d}t_1, \mathrm{d}t_2) = (\mu_1(\mathrm{d}t_1) \times \mathrm{d}t_2/2\pi) + (\mu_2(\mathrm{d}t_2) \times \mathrm{d}t_1/2\pi)$ in (4.1) where μ_1 and μ_2 are positive Borel measures on $[\pi, 3\pi/2)$. Then by mapping (4.1) to U^2 via $\zeta_j = (s_j - i)/(s_j + i)$, $e^{it_j} = (z_j - i)/(z_j + 1)$, $j = 1, 2$, and $f = iK$, and imposing $K(\infty, \infty) = 0$, we obtain

$$K(s_1, s_2) = K_1(s_1) + K_2(s_2), \tag{5.5}$$

where

$$K_1(s_1) = \int_0^1 \frac{\mu_1(\mathrm{d}z_1)}{s_1 - z_1}, \quad K_2(s_2) = \int_0^1 \frac{\mu_2(\mathrm{d}z_2)}{s_2 - z_2}. \tag{5.6}$$

In (5.6) $\mu_1(\mathrm{d}z_1)$ and $\mu_2(\mathrm{d}z_2)$ are new positive Borel measures on $[0, 1)$.

Now, for the Wiener bounds we assume that the volume fractions of the three materials p_1, p_2 and $p_3 = 1 - p_1 - p_2$ are known as well as $s_1, s_2 > 1$. In other words, F in (5.3) is known to first order,

$$F(s_1, s_2) = \frac{p_1}{s_1} + \frac{p_2}{s_2} + \ldots. \tag{5.7}$$

We now state

HYPOTHESIS 1. *If $K(s_1, s_2)$ in (5.4) is known to first order, i.e. if for fixed $\alpha_1, \alpha_2 > 0$*

$$K(s_1, s_2) = \frac{\alpha_1}{s_1} + \frac{\alpha_2}{s_2} + \ldots, \tag{5.8}$$

then for fixed $s_1, s_2 > 1$, K is minimized by (5.5), where K_1 and K_2 in (5.6) are separately minimized subject to

$$K_1(s_1) = \frac{\alpha_1}{s_1} + \ldots, \quad K_2(s_2) = \frac{\alpha_2}{s_2} + \ldots. \tag{5.9}$$

Note that in Hypothesis 1, K in (5.4) does not have the form (5.5) in general. The upshot of the conjecture is that K attains its minimum within the special class (5.5).

We apply this hypothesis to $K = F$ under (5.7). The minimum of (5.5) is obtained by letting $\mu_1 = p_1\delta_0$ and $\mu_2 = p_2\delta_0$ in (5.6). Then

$$F(s_1, s_2) \geqslant \frac{p_1}{s_1} + \frac{p_2}{s_2}, \tag{5.10}$$

which is an upper bound on ε^*. The lower bound is obtained by considering

$$H(t_1, t_2) = 1 - \frac{\varepsilon_3}{\varepsilon^*} = \frac{F(s_1, s_2)}{F(s_1, s_2) - 1}, \tag{5.11}$$

where $t_1 = 1 - s_1$ and $t_2 = 1 - s_2$. The function $H(t_1, t_2)$ has the same analyticity properties as $F(s_1, s_2)$ and has the first order expansion

$$H(t_1, t_2) = \frac{p_1}{t_1} + \frac{p_2}{t_2} + \ldots. \tag{5.12}$$

Applying Hypothesis 1 to $K = H$ gives

$$H(t_1, t_2) \geqslant \frac{p_1}{t_1} + \frac{p_2}{t_2}. \tag{5.13}$$

In terms of ε^*, (5.10) and (5.13) become

$$\left(1\Big/\frac{p_1}{\varepsilon_1} + \frac{p_2}{\varepsilon_2} + \frac{p_3}{\varepsilon_3}\right) \leqslant \varepsilon^* \leqslant p_1\varepsilon_1 + p_2\varepsilon_2 + p_3\varepsilon_3, \tag{5.14}$$

which are the classical Wiener bounds. They are optimal, and are attained by parallel plane configurations of the materials.

If the material is further assumed to be statistically isotropic, then the second order terms in (5.3) can be calculated (GOLDEN, 1984). Then F has the expansion

$$F(s_1,s_2)=\frac{p_1}{s_1}+\frac{p_2}{s_2}+\frac{p_1-p_1^2}{ds_1^2}+\frac{p_2-p_2^2}{ds_2^2}-\frac{2p_1p_2}{ds_1s_2}+\dots. \tag{5.15}$$

Note that there arises in (5.15) a nonzero second order cross term, $-2p_1p_2/ds_1s_2$, so that (5.5) is not applicable. However, consider the function

$$G(s_1,s_2)=\frac{F(s_1,s_2)}{1-\frac{1}{d}F(s_1,s_2)}, \tag{5.16}$$

which has the same analyticity properties as F so that its counterpart in D^2 has the representation (4.1). To second order $G(s_1,s_2)$ has the expansion

$$G(s_1,s_2)=\frac{p_1}{s_1}+\frac{p_2}{s_2}+\frac{p_1}{ds_1^2}+\frac{p_2}{ds_2^2}+\frac{0}{s_1s_2}+\dots. \tag{5.17}$$

The important point to note is that there is no second order cross term in (5.17), i.e. the transformation (5.16) has diagonalized the perturbation expansion (5.15) to second order. To obtain the bounds we now state

HYPOTHESIS 2. *If $K(s_1,s_2)$ in (5.4) is known to have the following expansion to second order for fixed $\alpha_1,\beta_1,\alpha_2,\beta_2>0$,*

$$K(s_1,s_2)=\frac{\alpha_1}{s_1}+\frac{\alpha_2}{s_2}+\frac{\beta_1}{s_1^2}+\frac{\beta_2}{s_2^2}+\frac{0}{s_1s_2}+\dots, \tag{5.18}$$

then for fixed $s_1,s_2>1$, K is minimized by (5.5), where K_1 and K_2 in (5.6) are separately minimized subject to

$$K_1(s_1)=\frac{\alpha_1}{s_1}+\frac{\beta_1}{s_1^2}+\dots,\quad K_2(s_2)=\frac{\alpha_2}{s_2}+\frac{\beta_2}{s_2^2}+\dots. \tag{5.19}$$

We apply Hypothesis 2 to $K=G$ and call $K_1=G_1$ and $K_2=G_2$, which are both known to second order,

$$G_1(s_1)=\frac{p_1}{s_1}+\frac{p_1}{ds_1^2}+\dots,\quad G_2(s_2)=\frac{p_2}{s_2}+\frac{p_2}{ds_2^2}+\dots. \tag{5.20}$$

To incorporate the constraints (5.20) into the minimization of G_1 and G_2, we use the transformation procedure developed for two-component media. Let

$$J_1=\frac{1}{p_1}-\frac{1}{s_1G_1}, \tag{5.21}$$

and similarly for G_2. Then to first order

$$J_1(s_1)=\frac{1/dp_1}{s_1}+\dots. \tag{5.22}$$

Since $J_1(s_1)$ has the same type of integral representation as does $K_1(s_1)$ in (5.6), it is minimized for fixed $s_1>1$ by the measure $(1/dp_1)\delta_0$, so that

$$J_1(s_1)\geqslant\frac{1/dp_1}{s_1}. \tag{5.23}$$

Combining (5.23) with the analogous result for G_2 yields for G,

$$G(s_1,s_2)\geqslant\frac{p_1}{s_1-\frac{1}{d}}+\frac{p_2}{s_2-\frac{1}{d}}, \tag{5.24}$$

which is an upper bound on ε^*. To get the lower bound for $\varepsilon_1\leqslant\varepsilon_2\leqslant\varepsilon_3$ we consider $\tilde F(q_2,q_3)=1-\varepsilon^*/\varepsilon_1$, where $q_2=1/(1-\varepsilon_2/\varepsilon_1)$ and $q_3=1/(1-\varepsilon_3/\varepsilon_1)$. The function $\tilde F(q_2,q_3)$ has the same analyticity properties as $F(s_1,s_2)$ and to second order has the expansion

$$\tilde F(q_2,q_3)=\frac{p_2}{q_2}+\frac{p_3}{q_3}+\frac{p_2-p_2^2}{dq_2^2}+\frac{p_3-p_3^2}{dq_3^2}-\frac{2p_1p_2}{dq_2q_3}+\dots. \tag{5.25}$$

As in the above treatment for the upper bound on ε^*, we consider

$$\tilde G(q_2,q_3)=\frac{\tilde F(q_2,q_3)}{1-\frac{1}{d}\tilde F(q_2,q_3)}, \tag{5.26}$$

which again diagonalizes to second order the expansion (5.25). Under Hypothesis 2 with $K=\tilde G$, we have

$$\tilde G(q_2,q_3)\leqslant\frac{p_2}{q_2-\frac{1}{d}}+\frac{p_3}{q_3-\frac{1}{d}}, \tag{5.27}$$

where now we get an upper bound because $q_2,q_3<0$ when $\varepsilon_1\leqslant\varepsilon_2\leqslant\varepsilon_3$. In terms of ε^*, (5.24) and (5.27) become

$$\varepsilon_1+1\Big/\left(\frac{1}{A_1}-\frac{1}{d\varepsilon_1}\right)\leqslant\varepsilon^*\leqslant\varepsilon_3+1\Big/\left(\frac{1}{A_3}-\frac{1}{d\varepsilon_3}\right),\quad A_j=\sum_{l=1}^{3}p_l\Big/\left(\frac{1}{\varepsilon_l-\varepsilon_j}+\frac{1}{d\varepsilon_j}\right), \tag{5.28}$$

which are the Hashin-Shtrikman bounds for three-component materials. These bounds are optimal for certain volume fraction regimes (MILTON, 1981d). The upper bound is attained by a mixture of all different sized spheres of ε_1 and ε_2 each coated with ε_3 in the appropriate volume fraction. The lower bound is attained by the same mixture with ε_1 and ε_3 reversed.

We remark here that the above arguments are immediately generalizable to N-component materials. The function $K(s_1,\dots,s_n)$, $n=N-1$, analogous to (5.4), is decomposed as in (5.5) into a sum of n pieces, $K_1(s_1),\dots,K_n(s_n)$. The extremization procedure is then carried out for each piece. The only point that needs checking is that the G transformation in (5.16) diagonalizes to second order the perturbation expansion of F for N-component materials. A simple calculation shows that it does.

6. New Complex Bounds for Three-Component Media

We first give bounds on ε^* assuming no information about the material aside from ε_1, ε_2 and ε_3 which are now complex. These bounds are the analogues of (3.7) for three-component media and are determined primarily by the condition $F(1,1) \leqslant 1$. They were first shown to the author by MILTON (1984) in a form to be described later, but can be derived in the context of this paper by using

HYPOTHESIS 3. *Suppose $K(s_1, s_2)$ in (5.4) is subjected only to $K(1,1) \leqslant 1$ for fixed $(s_1, s_2) \in U^2$. Then boundary points of the allowed region for K can be obtained by first setting $\mu_2 = 0$ in (5.5) and (5.6) and then extremizing K_1 subject to $K_1(1) \leqslant 1$, and then setting $\mu_1 = 0$ in (5.5) and (5.6) and extremizing K_2 subject to $K_2(1) \leqslant 1$.*

In particular, when extremizing K_1 we let $\mu_1 = \alpha_1 \delta_{z_1}$ in (5.6) so that

$$K_1(s_1) = \frac{\alpha_1}{s_1 - z_1}, \tag{6.1}$$

and the condition that $K_1(1) \leqslant 1$ becomes $\alpha_1 + z_1 \leqslant 1$. Then the image of either $\{(\alpha_1, z_1): z_1 = 0,\ 0 \leqslant \alpha_1 \leqslant 1\}$ or $\{(\alpha_1, z_1): \alpha_1 + z_1 = 1,\ \alpha_1, z_1 \geqslant 0\}$ lies on the boundary of the allowed region for K_1, and similarly for K_2. Note that we do not generate the entire boundary of the allowed region using this procedure applied only to K.

We now obtain the above described zeroth order bounds by applying the above hypothesis to $K = F$. If $\mu_2 = 0$ in (5.6), then one of the following two arcs joining 0 and $1/s_1$ lies on the boundary of the allowed region in the F-plane,

$$L_1(\alpha_1) = \frac{\alpha_1}{s_1}, \quad C_1(\alpha_1) = \frac{\alpha_1}{s_1 - (1-\alpha_1)}, \quad 0 \leqslant \alpha_1 \leqslant 1. \tag{6.2}$$

If $\mu_1 = 0$ in (5.6), then one of the following two arcs joining 0 and $1/s_2$ lies on the boundary of the allowed region,

$$L_2(\alpha_2) = \frac{\alpha_2}{s_2}, \quad C_2(\alpha_2) = \frac{\alpha_2}{s_2 - (1-\alpha_2)}, \quad 0 \leqslant \alpha_2 \leqslant 1. \tag{6.3}$$

Hypothesis 3 can then be applied to $\tilde{F}(q_2, q_3) = 1 - \varepsilon^*/\varepsilon_1$ to obtain a line segment and a circular arc, each of which joins $1/s_1$ and $1/s_2$, thus filling out the boundary of the allowed region. The other segment and circular arc obtained from $\tilde{F}(q_2, q_3)$ will be the same as one of the above pairs from $F(s_1, s_2)$.

In the F-plane the allowed region may be described as follows. Let T_L be the triangular region lying inside the three line segments above with vertices 0, $1/s_1$, and $1/s_2$. Let T_C be the curvilinear triangular region lying inside the three circular arcs above with vertices 0, $1/s_1$ and $1/s_2$. Then the allowed region is the union of T_L and T_C.

MILTON (1984) has pointed out to the author that these bounds have a simple interpretation in the ε^*-plane and are easily shown to be optimal. Let T_L^* be the triangular region bounded by the line segments

$$\begin{aligned} L_{12}(\alpha) &= \alpha\varepsilon_1 + (1-\alpha)\varepsilon_2, \\ L_{23}(\alpha) &= \alpha\varepsilon_2 + (1-\alpha)\varepsilon_3, \\ L_{13}(\alpha) &= \alpha\varepsilon_1 + (1-\alpha)\varepsilon_3, \quad 0 \leqslant \alpha \leqslant 1. \end{aligned} \tag{6.4}$$

Let T_C^* be the curved triangular region bounded by the circular arcs

$$\begin{aligned} C_{12}(\alpha) &= 1/(\alpha/\varepsilon_1 + (1-\alpha)/\varepsilon_2), \\ C_{23}(\alpha) &= 1/(\alpha/\varepsilon_2 + (1-\alpha)/\varepsilon_3), \\ C_{13}(\alpha) &= 1/(\alpha/\varepsilon_1 + (1-\alpha)/\varepsilon_3), \quad 0 \leqslant \alpha \leqslant 1. \end{aligned} \tag{6.5}$$

Each circular arc $C_{ij}(\alpha)$, when extended, passes through ε_i, ε_j, and the origin. The allowed region is again the union of T_L^* and T_C^*. All segments and arcs are attainable. The line segment L_{ij} is attained by a slab geometry parallel to the applied field composed of materials i and j in the volume fractions α and $1-\alpha$. The circular arc C_{ij} is attained by the same slab geometry but arranged perpendicular to the field. The outermost arcs of the zeroth order bounds for particular values of the component dielectric constants are depicted as arcs a, b, and c in Fig. 1.

If the volume fractions p_1, p_2 and p_3 are known as well as ε_1, ε_2 and ε_3, then F is known to first order as in (5.7). To get complex bounds on ε^* which incorporate this information, we state

HYPOTHESIS 4. *If $K(s_1, s_2)$ in (5.4) is known to first order, i.e. if for fixed $\alpha_1, \alpha_2 > 0$*

$$K(s_1, s_2) = \frac{\alpha_1}{s_1} + \frac{\alpha_2}{s_2} + \ldots, \tag{6.6}$$

then for fixed $(s_1, s_2) \in U^2$ the values of $K(s_1, s_2)$ lie inside the region generated by the sum of the two circles that contain the values of $K_1(s_1)$ and $K_2(s_2)$ in (5.6) subject to

$$K_1(s_1) = \frac{\alpha_1}{s_1} + \ldots, \quad K_2(s_2) = \frac{\alpha_2}{s_2} + \ldots. \tag{6.7}$$

With K_1 as in (6.7), we obtain its extremal values by letting $\mu_1 = \alpha_1 \delta_{z_1}$ in (5.6). The allowed values of K_1 lie inside the circle

$$K_1(s_1) = \frac{\alpha_1}{s_1 - z_1}, \quad -\infty \leqslant z_1 \leqslant \infty. \tag{6.8}$$

Applying Hypothesis 4 to $K = F$ under (5.7) then restricts the values of $F = 1 - \varepsilon^*/\varepsilon_3$ to lie inside the region R_3 generated by

$$F(s_1, s_2) = \frac{p_1}{s_1 - z_1} + \frac{p_2}{s_2 - z_2}, \quad -\infty \leqslant z_1, z_2 \leqslant \infty. \tag{6.9}$$

Notice that we have ignored the support restrictions on μ_1 and μ_2 inherent in (5.6), so that these bounds are seemingly very crude. We will see later, though, that under certain circumstances these bounds are optimal.

The region R_3 can be constructed as follows. First, $p_1/(s_1-z_1)$ and $p_2/(s_2-z_2)$ with $-\infty \leqslant z_1, z_2 \leqslant \infty$ are both circles in the lower half plane that contain the origin and are thus tangential to the real axis. Now add to each point $p_1/(s_1-z_1)$ the circle $p_2/(s_2-z_2)$, $-\infty \leqslant z_2 \leqslant \infty$. The outside boundary of R_3 is a circle characterized by

$$\arg\left(\frac{\partial F}{\partial z_1}\right) = \arg\left(\frac{\partial F}{\partial z_2}\right), \tag{6.10}$$

where "arg" denotes argument and F is as in (6.9). This condition is equivalent to $\arg(s_1-z_1) = \arg(s_2-z_2)$, or,

$$z_2 = \frac{b_2}{b_1}\, z_1 + \left(a_2 - \frac{b_2 a_1}{b_1}\right), \tag{6.11}$$

where $s_1 = a_1 + ib_1$ and $s_2 = a_2 + ib_2$. Thus the outside boundary $B_3(z_1)$ of R_3 is the image of the line (6.11) in (z_1, z_2)-space under (6.9) and can be parameterized in the F-plane by

$$B_3(z_1) = \frac{p_1 + p_2 b_1/b_2}{s_1 - z_1}, \quad -\infty \leqslant z_1 \leqslant \infty. \tag{6.12}$$

With $\arg \varepsilon_3 < \arg \varepsilon_2 < \arg \varepsilon_1$, we have that $(q_2, q_3) \in L^2$, where $L^2 = \{\text{Im } q_2 < 0\} \times \{\text{Im } q_3 < 0\}$. Then we can apply Hypothesis 4 to $\tilde{F}(q_2, q_3)$, which restricts its values to a circular region R_1 bounded by

$$B_1(z_1) = \frac{p_2 + p_3 \text{ Im } q_2/\text{Im } q_3}{q_2 - z_1}, \quad -\infty \leqslant z_1 \leqslant \infty. \tag{6.13}$$

In the ε^*-plane the regions R_1 and R_3 become circular regions R_1^* and R_3^*. Then, a complex extension of the Wiener bounds (5.14) is the intersection of $(R_1^* \cap R_3^*)$ with the zeroth order complex bounds described above. The two arcs arising from (6.12) and (6.13) in the ε^*-plane are depicted as arcs d and e in Fig. 1.

In the above bounds we have not considered the function $1-\varepsilon^*/\varepsilon_2$ because with $\arg \varepsilon_3 < \arg \varepsilon_2 < \arg \varepsilon_1$, $1/(1-\varepsilon_1/\varepsilon_2)$ is in the upper half plane while $1/(1-\varepsilon_3/\varepsilon_2)$ is in the lower half plane. We have so far only considered functions on U^2 or L^2. Better bounds can be obtained by considering $1-\varepsilon^*/\varepsilon_2$ on the domain of analyticity given in Sect. 2, which is more general than U^2 or L^2. Furthermore, we have not included the functions $1-\varepsilon_3/\varepsilon^*$ and $1-\varepsilon_1/\varepsilon^*$, because they give the same first order bounds on ε^* as do $1-\varepsilon^*/\varepsilon_3$ and $1-\varepsilon^*/\varepsilon_1$, as a simple calculation shows.

We now discuss the optimality of the new complex Wiener bounds. Let us focus on $F = 1-\varepsilon^*/\varepsilon_3$. Recall that for an actual material, z_1 and z_2 in (6.9) are restricted by $0 \leqslant z_1, z_2 < 1$ and $F(1,1) \leqslant 1$. These conditions define a region Z in the lower left hand corner of the unit square in (z_1, z_2)-space. This region Z is bounded by the line segments $(0,0) \to (p_3/(1-p_2), 0)$ and $(0,0) \to (0, p_3/(1-p_1))$, and the hyperbolic arc defined by $p_1/(1-z_1) + p_2/(1-z_2) = 1$, $0 \leqslant z_1, z_2 < 1$. If the line (6.11) passes through the region Z in (z_1, z_2)-space then the following geometry (MILTON, 1984) attains that section of $B_3(z_1)$. Expression (6.9) can be written as

$$\varepsilon^* = \left(1 - \frac{p_1}{1-z_1} - \frac{p_2}{1-z_2}\right)\varepsilon_3 + \left(\frac{p_1}{1-z_1}\right)\left(1\Big/\left(\frac{1-z_1}{\varepsilon_1} + \frac{z_1}{\varepsilon_3}\right)\right) + \left(\frac{p_2}{1-z_2}\right)\left(1\Big/\left(\frac{1-z_2}{\varepsilon_2} + \frac{z_2}{\varepsilon_3}\right)\right). \tag{6.14}$$

This expression is clearly the effective dielectric constant of a composite consisting of slabs of three materials parallel to the field in the volume fractions $(1-p_1/(1-z_1)-p_2/(1-z_2))$, $(p_1/(1-z_1))$, and $(p_2/(1-z_2))$. The first material is just ε_3. The second is a slab composite perpendicular to the field composed of ε_1 in the volume fraction $1-z_1$ and ε_3 in the volume fraction z_1. The third is a slab composite perpendicular to the field composed of ε_2 in the volume fraction $1-z_2$ and ε_3 in the volume fraction z_2. The attained arc is traced out as z_1 varies so that (z_1, z_2) stays in Z according to (6.11).

We will now show that an arc of $B_3(z_1)$ different from but possibly overlapping that above can be attained by an actual material. The circle generated by $H(t_1, t_2) = 1-\varepsilon_3/\varepsilon^*$ is the same as that generated by $F(s_1, s_2) = 1-\varepsilon^*/\varepsilon_3$, and can be parameterized in the H-plane by

$$H(t_1, t_2) = \frac{p_1}{t_1 - z_1} + \frac{p_2}{t_2 - z_2}, \tag{6.15}$$

where, analogous to (6.11), we have

$$z_2 = \frac{b_2}{b_1}\, z_1 + \left(1 - a_2 - \frac{b_2}{b_1}\,(1-a_1)\right). \tag{6.16}$$

Now the admissible region analogous to Z in (z_1, z_2)-space for H certainly lies in the unit square. In terms of H, the condition $F(1,1) \leqslant 1$ translates into $H(0,0) \leqslant 0$, which is automatically satisfied since $0 \leqslant z_1, z_2 \leqslant 1$. However, the condition that $F(0,0) \leqslant 0$ translates into $H(1,1) \leqslant 1$, so that the admissible region in (z_1, z_2)-space for H is identical to Z. When the line (6.16) passes through Z, the section of $B_3(z_1)$ that corresponds to the line segment in Z is attained by the following geometry. Analogous to (6.14), (6.15) may be written as

$$\varepsilon^* = 1\Big/\left(\left[\left(1 - \frac{p_1}{1-z_1} - \frac{p_2}{1-z_2}\right)\Big/\varepsilon_3\right] + \left[\left(\frac{p_1}{1-z_1}\right)\Big/((1-z_1)\varepsilon_1 + z_1\varepsilon_3)\right] + \left[\left(\frac{p_2}{1-z_2}\right)\Big/((1-z_2)\varepsilon_2 + z_2\varepsilon_3)\right]\right). \tag{6.17}$$

The material corresponding to (6.17) is the composite in (6.14) rotated by 90°. A similar analysis to that above can be given to show that parts of the arc $B_1(z_1)$ can be attained by the same type of slab geometry but with the host material being ε_1. We remark here that since only parts of our bounds are attainable under certain circumstances, we expect that the bounds can be improved. (In fact, subsequent to this work, BERGMAN (1985) has improved these bounds by finding arcs analogous to d and e which are in the interior of the region enclosed by arcs a, b, and c.)

An interesting feature that seems to distinguish the complex Wiener bounds for three-component media from those for two-component media is as follows. For two-component media, the slab geometries corresponding to $\varepsilon^* = p_1\varepsilon_1+p_2\varepsilon_2$ and $\varepsilon^* = 1/(p_1\varepsilon_1+p_2\varepsilon_2)$ have values for ε^* that form the vertices of the complex bound. For three-component media, however, $\varepsilon^* = p_1\varepsilon_1+p_2\varepsilon_2+p_3\varepsilon_3$ and $\varepsilon^* = 1/(p_1/\varepsilon_1+p_2/\varepsilon_2+p_3/\varepsilon_3)$, which are depicted as points A and B in Fig. 1, lie inside the complex bound. Only in the very special case when the line (6.11) passes through the origin does point A lie on the bound, and only when (6.16) passes through the origin does point B lie on the bound. Even though our new bounds in their present form do not reduce to the bounds (5.14) when the parameters become real, we still call them complex extensions of the Wiener bounds because they rely on the same amount of information, namely (5.7).

If we now further assume that the material is statistically isotropic, then F is known to second order as in (5.15). The function G then has a diagonal expansion to second order,

$$G(s_1,s_2) = \frac{p_1}{s_1}+\frac{p_2}{s_2}+\frac{p_1}{ds_1^2}+\frac{p_2}{ds_2^2}+\frac{0}{s_1s_2}+\dots. \tag{6.18}$$

To obtain complex versions of the Hashin–Shtrikman bounds we state

HYPOTHESIS 5. *If $K(s_1,s_2)$ in (5.4) is known to have the following second order expansion for fixed $\alpha_1,\alpha_2,\beta_1,\beta_2>0$,*

$$K(s_1,s_2) = \frac{\alpha_1}{s_1}+\frac{\alpha_2}{s_2}+\frac{\beta_1}{s_1^2}+\frac{\beta_2}{s_2^2}+\frac{0}{s_1s_2}+\dots, \tag{6.19}$$

then for fixed $(s_1,s_2)\in U^2$ the values of $K(s_1,s_2)$ lie inside the region generated by the sum of the two circles that contain the values of $K_1(s_1)$ and $K_2(s_2)$ in (5.6) subject to

$$K_1(s_1) = \frac{\alpha_1}{s_1}+\frac{\beta_1}{s_1^2}+\dots,\quad K_2(s_2) = \frac{\alpha_2}{s_2}+\frac{\beta_2}{s_2^2}+\dots. \tag{6.20}$$

To incorporate the constraints (6.20) into $K_1 = G_1$ and $K_2 = G_2$ we again use the transformation procedure developed for a single complex variable. With J_1 as in (5.21) and (5.22), its values lie inside the circle

$$J_1(s_1) = \frac{1/dp_1}{s_1-z_1},\quad -\infty \leqslant z_1 \leqslant \infty. \tag{6.21}$$

Equivalently, the values of $G_1(s_1)$ lie inside the circle

$$G_1(s_1) = \frac{p_1(s_1-z_1)}{s_1(s_1-z_1-1/d)},\quad -\infty \leqslant z_1 \leqslant \infty. \tag{6.22}$$

Then the values of G lie inside the region R_G generated by

$$G(s_1,s_2) = \frac{p_1(s_1-z_1)}{s_1(s_1-z_1-1/d)}+\frac{p_2(s_2-z_2)}{s_2(s_2-z_2-1/d)},\quad -\infty \leqslant z_1,z_2 \leqslant \infty. \tag{6.23}$$

The outer circular boundary of R_G is obtained by setting

$$\arg\left(\frac{\partial G}{\partial z_1}\right) = \arg\left(\frac{\partial G}{\partial z_2}\right), \tag{6.24}$$

which implies that

$$\tan^{-1}\left(\frac{b_2}{a_2-z_2-1/d}\right) = \tan^{-1}\left(\frac{b_1}{a_1-z_1-1/d}\right)+\frac{1}{2}\left[\tan^{-1}\left(\frac{b_1}{a_1}\right)-\tan^{-1}\left(\frac{b_2}{a_2}\right)\right]. \tag{6.25}$$

Using the rule for the tangent of a sum gives after some manipulation

$$z_2 = a_2-1/d+\frac{b_2(cb_1-a_1+1/d)+b_2z_1}{(b_1+ca_1-c/d)-cz_1}, \tag{6.26}$$

where

$$c = \tan\frac{1}{2}\left[\tan^{-1}\left(\frac{b_1}{a_1}\right)-\tan^{-1}\left(\frac{b_2}{a_2}\right)\right]. \tag{6.27}$$

Hypothesis 5 can be applied to $\tilde{G}(q_2,q_3)$, which restricts its values to a circular region $R_{\tilde{G}}$ in the $\tilde{G}$-plane. Then a similar analysis to that above is applied to get the outer circular boundary.

Next we consider $H(t_1,t_2) = 1-\varepsilon_3/\varepsilon^*$ which to second order gives a different bound than $F = 1-\varepsilon^*/\varepsilon_3$. The expansion of H to second order is

$$H(t_1,t_2) = \frac{p_1}{t_1}+\frac{p_2}{t_2}+\frac{((d-1)/d)(p_1-p_1^2)}{t_1^2}+\frac{((d-1)/d)(p_2-p_2^2)}{t_2^2} - \frac{2((d-1)/d)p_1p_2}{t_1t_2}+\dots. \tag{6.28}$$

Then the transformation

$$L(t_1,t_2) = \frac{H}{1-((d-1)/d)H} \tag{6.29}$$

has to second order the expansion

$$L(t_1,t_2) = \frac{p_1}{t_1}+\frac{p_2}{t_2}+\frac{((d-1)/d)p_1}{t_1^2}+\frac{((d-1)/d)p_2}{t_2^2}+\frac{0}{t_1t_2}+\dots. \tag{6.30}$$

Hypothesis 5 can then be applied to $L(t_1,t_2)$, yielding a region R_L in the L-plane. Finally we consider $\tilde{H}(v_2,v_3) = 1-\varepsilon_1/\varepsilon^*$, where $v_2 = 1/(1-\varepsilon_1/\varepsilon_2)$ and $v_3 = 1/(1-\varepsilon_1/\varepsilon_3)$. Hypothesis 5 can be applied to $\tilde{L}$, which is the analogue of (6.29) for $\tilde{H}$, yielding a region $R_{\tilde{L}}$ in the $\tilde{L}$-plane. In the ε^*-plane, the four regions R_G, $R_{\tilde{G}}$, R_L and $R_{\tilde{L}}$ become the four circular regions R_G^*, $R_{\tilde{G}}^*$, R_L^* and $R_{\tilde{L}}^*$. A complex version of the Hashin–Shtrikman bounds (5.28) is then the intersection of $(R_G^*\cap R_{\tilde{G}}^*\cap R_L^*\cap R_{\tilde{L}}^*)$ with the complex Wiener bounds described above. In Fig. 1 this region is enclosed by arcs f and g. Arc f comes from the circular boundary of R_G^* and arc g comes from the circular boundary of $R_{\tilde{G}}^*$. For the values of the parameters used in Fig. 1, the bounds induced by L and $\tilde{L}$ are outside those induced by G and $\tilde{G}$, and have therefore been omitted.

Note that when $z_1 = z_2 = 0$ in (6.23), $\varepsilon^* = \varepsilon_3+1/(1/A_3-1/d\varepsilon_3)$ as in (5.28), which is the permittivity of an actual composite, as mentioned before. Only in the special case when the curve defined by (6.26) runs through $(z_1,z_2) = (0,0)$ does the permittivity of this composite lie on the circle which bounds R_G. A similar statement can be made for $\varepsilon^* = \varepsilon_1 + 1/(1/A_1 - 1/d\varepsilon_1)$ and $R_{\tilde{G}}$. These two points, denoted by C and D in Fig. 1, lie inside the complex Hashin-Shtrikman bounds in Fig. 1. An indication that these bounds are quite crude is that points A and B lie inside them. Presumably these second order bounds can be improved.

ACKNOWLEDGEMENTS

The author is very pleased to thank G. Papanicolaou for his enthusiasm, encouragement, and guidance during the course of the work. In addition, the author would like to acknowledge useful conversations with R. Caflisch, R. Kohn, P. Lax, and A. Sokal. Special thanks are owed to G. Milton for his help through many valuable discussions and letters. This work was primarily done while the author held a Hertz Foundation Graduate Fellowship. The research was supported in part by the Air Force Office of Scientific Research through Grant AFOSR-80-0228, and the author thanks both of these sources. The manuscript was completed while the author was a visitor at Institut des Hautes Etudes Scientifiques, and he thanks N. Kuiper for the hospitality.

REFERENCES

AKHIEZER, N. I. and GLAZMAN, I. M. 1966 *The Theory of Linear Operators in Hilbert Space*, F. Ungar, New York.
BAKER, G. A. Jr. 1969 *J. Math. Phys.* **10,** 814.
BAKER, G. A. Jr. 1975 *Essentials of Padé Approximants*, Academic Press, New York.
BERGMAN, D. J. 1978 *Phys. Rep.* **C43,** 377.
BERGMAN, D. J. 1979 *J. Phys.* **C12,** 4947.
BERGMAN, D. J. 1980a *Phys. Rev.* **B22,** 3527.
BERGMAN, D. J. 1980b *Phys. Rev. Lett.* **44,** 1285.
BERGMAN, D. J. 1981 *Phys. Rev.* **B23,** 3058.
BERGMAN, D. J. 1982a *Lecture Notes in Physics*, Vol. 154, p. 10, Springer, New York.
BERGMAN, D. J. 1982b *Ann. Phys.* **138,** 78.
BERGMAN, D. J. 1983 *Bulk Physical Properties of Composite Media*, Lecture Notes, L'Ecole d'Eté d'Analyse Numérique.
BERGMAN, D. J. 1985 Proceedings of the U. of Minnesota Workshop on Homogenization and Effective Moduli, Minneapolis, Minn.
DUNFORD, N. and SCHWARTZ, J. T. 1958 *Linear Operators, Part I*, Interscience, New York.
FELDERHOF, B. U. 1984 *Physica* **126A,** 430.
GOLDEN, K. 1984 Ph.D. Thesis, New York University, New York.
GOLDEN, K. and PAPANICOLAOU, G. 1983 *Comm. Math. Phys.* **90,** 473.
GOLDEN, K. and PAPANICOLAOU, G. 1985 *J. Stat. Phys.* **40,** 655.
HASHIN, Z. and SHTRIKMAN, S. 1962 *Appl. Phys.* **33,** 3125.
KANTOR, Y. and BERGMAN, D. J. 1984 *J. Mech. Phys. Solids* **32,** 41.
KELLER, J. B. 1964 *J. Math. Phys.* **5,** 548.
KOHLER, W. and PAPANICOLAOU, G. 1982 *Lecture Notes in Physics*, Vol. 154, p. 111, Springer, New York

LURIE, K. A. and CHERKAEV, A. V. 1984 *Proc. Roy. Soc. Edinb.* **99A,** 71.
MCDONALD, J. N. 1982 *Michigan Math. J.* **29,** 259.
MENDELSON, K. S. 1975 *J. Appl. Phys.* **46,** 4740.
MILTON, G. W. 1980 *Appl. Phys. Lett.* **37,** 300.
MILTON, G. W. 1981a *J. Appl. Phys.* **52,** 5294.
MILTON, G. W. 1981b *J. Appl. Phys.* **52,** 5286.
MILTON, G. W. 1981c *Phys. Rev. Lett.* **46,** 542.
MILTON, G. W. 1981d *Appl. Phys.* **A26,** 125.
MILTON, G. W. 1982 *J. Mech. Phys. Solids* **30,** 177.
MILTON, G. W. 1984 Private communication.
MILTON, G. W. and GOLDEN, K. 1985 *Thermal Conductivity 18*, Eds. T. Ashworth and D. R. Smith, p. 571. Plenum, New York.
PAPANICOLAOU, G. and VARADHAN, S. 1982 *Colloquia Mathematica Societatis János Bolyai 27, Random Fields*, p. 835. North Holland, Amsterdam.
RUDIN, W. 1970 *Actes, Congrès Intern. Math.*, p. 489.
RUDIN, W. 1983 Private communication.
SCHULGASSER, K. 1976 *J. Math. Phys.* **17,** 378.
TARTAR, L. and MURAT, F. 1981 Private communication (through R. Kohn).
WALL, H. S. 1948 *Analytic Theory of Continued Fractions*, D. Van Nostrand, New York.
WIENER, O. 1912 *Abhandl. Math.-Phys. Kl. Königl. Sachsischen Ges.* **32,** 509.

APPENDIX

HIGHER ORDER BOUNDS FOR TWO-COMPONENT MEDIA

If $F(s)$ in Sect. 3 is known to nth order via (3.4), then the transformation (3.13) can be iterated to obtain bounds on ε^* which incorporate this information. BAKER (1969, 1975) was the first to use such an iteration procedure in a slightly different form to obtain nth order complex bounds on half plane functions like F. His work was done in the context of Padé approximants to Stieltjes series. Independently, in the context of heterogeneous media MILTON (1981a) used another method to obtain nth order complex bounds on ε^*, which reduce for real component parameters to those obtained from variational principles. FELDERHOF (1984) gives another formulation of Milton's nth order bounds. Baker's, Milton's, Felderhof's, and the following bounds all solve essentially the same problem, except that Baker's are more general in the sense that his were derived without reference to a particular physical problem. The relationship of Milton's bounds to bounds on Stieltjes series is discussed by MILTON and GOLDEN (1985), where a slightly different formulation of the iteration is given. BERGMAN (1982b, 1983) introduced the transformation (3.13) into the theory of heterogeneous media and it is in fact equivalent to Baker's transformation. He used it and a variant to derive second order bounds on ε^*. Subsequently, KANTOR and BERGMAN (1984) suggested iteration of (3.13).

To demonstrate the iteration, we derive third order bounds on ε^*. Assume

$$F(s) = \frac{a_1}{s} + \frac{a_2}{s^2} + \frac{a_3}{s^3} + \dots. \tag{A.1}$$

The first iterate F_1 in (3.13) has the representation

$$F_1(s) = \int_0^1 \frac{\mu_1(dz)}{s-z}, \tag{A.2}$$

where μ_1 is a positive measure on [0,1]. Under (A.1), F_1 is known to second order

$$F_1(s) = \frac{a_2}{a_1^2 s} + \left(\frac{a_3}{a_1^2} - \frac{a_2^2}{a_1^3}\right)\frac{1}{s^2} + \dots. \tag{A.3}$$

Now let

$$F_2(s) = \frac{a_1^2}{a_2} - \frac{1}{sF_1(s)} \tag{A.4}$$

so that

$$F_2(s) = \int_0^1 \frac{\mu_2\,(\mathrm{d}z)}{s-z}. \tag{A.5}$$

Then to first order,

$$F_2(s) = a_1\left(\frac{a_1 a_3}{a_2^2} - 1\right)\frac{1}{s} + \dots, \tag{A.6}$$

which forces the mass of μ_2 to be $a_1(a_1a_3/a_2^2 - 1)$. The considerations of Sect. 3 tell us that the values of $F_2(s)$ for fixed $s \in C$ lie inside the circle $a_1(a_1a_3/a_2^2-1)/(s-z)$, $-\infty \leqslant z \leqslant \infty$. Another circle is obtained by applying the same considerations to $E_2(s)$, the analog of (A.4) for $E(s)$. The vertices of the induced bounds in the ε^*-plane correspond to $z = 0$ on the F_2 circle and $z = 0$ on the E_2 circle, and lie on the arcs induced by (3.16) and (3.18). MILTON (1981a) shows that these bounds are optimal. In general the materials that attain them are anisotropic.

When ε_1 and ε_2 are real and positive, the above region collapses to an interval, where the upper bound is

$$\varepsilon^* \leqslant \varepsilon_2 + \left\{a_1\Big/\left(\frac{1}{\varepsilon_1-\varepsilon_2} + \frac{a_2/a_1}{\varepsilon_2 + [(a_2/a_1)-(a_3/a_2)](\varepsilon_2-\varepsilon_1)}\right)\right\}, \quad \varepsilon_1 \leqslant \varepsilon_2. \tag{A.7}$$

The lower bound is

$$\frac{1}{\varepsilon^*} \leqslant \frac{1}{\varepsilon_1} + e_1\Big/\left[1\Big/\left(\frac{1}{\varepsilon_2}-\frac{1}{\varepsilon_1}\right) + \frac{e_2}{e_1}\Big/\left(\frac{e_2}{e_1}-\frac{e_3}{e_2}\right)\left(\frac{1}{\varepsilon_1}-\frac{1}{\varepsilon_2}\right)\right], \quad \varepsilon_1 \leqslant \varepsilon_2, \tag{A.8}$$

where

$$E(s) = \frac{e_1}{s} + \frac{e_2}{s^2} + \frac{e_3}{s^3} + \dots. \tag{A.9}$$

Note that (A.7) with $a_1 = p_1$ and $a_2 = p_1p_2/d$ is a tighter bound than (3.20) since $(a_2/a_1 - a_3/a_2) \leqslant 0$, which is the Schwartz inequality, and similarly for (A.8).

The higher order bounds may be described as follows. Assume

$$F(s) = \frac{a_1}{s} + \frac{a_2}{s^2} + \dots + \frac{a_n}{s^n} + \dots. \tag{A.10}$$

Following the iteration procedure indicated above, the last transformation to be used will be

$$F_{n-1}(s) = \frac{1}{\mu_{n-2}^{(0)}} - \frac{1}{sF_{n-2}(s)}, \tag{A.11}$$

with

$$F_{n-1}(s) = \int_0^1 \frac{\mu_{n-1}\,(\mathrm{d}z)}{s-z}. \tag{A.12}$$

Then the values of $F_{n-1}(s)$ for fixed $s \in C$ lie inside the circle $\mu_{n-1}^{(0)}/(s-z)$, $-\infty \leqslant z \leqslant \infty$. Since

F is fractional linear in F_{n-1}, this circle transforms into a circle in the F-plane,

$$C_n(z) = \left\langle 1\Big/\frac{s}{\mu^{(0)}} - \left\{s\Big/\frac{s}{\mu_1^{(0)}} - \left[s\Big/\frac{s}{\mu_2^{(0)}} - \left[s\Big/ \dots \Big/\left(\frac{s}{\mu_{n-2}^{(0)}} - \frac{s\mu_{n-1}^{(0)}}{s-z}\right)\right]\right]\right\}\right\rangle. \tag{A.13}$$

The $\mu_i^{(0)}$ can be computed from the a_i as the iteration proceeds. Within the theory of Padé approximants there is a literature which covers the determination of such quantities as $\mu_i^{(0)}$ (BAKER, 1969, 1975; WALL, 1948). Analysis of E_{n-1}, the analogue of (A.11) for $E(s)$, gives a circle $\hat{C}_n(z)$ in the E-plane analogous to (A.13). The intersection of the two induced regions in the ε^* plane is bounded by two circular arcs. The arc coming from (A.13) is traced out as z varies between 0 and a, where a is determined by the condition that $F(1) \leqslant 1$, just as in the first and second order bounds, and similarly for E. The resulting bounds form a nested sequence of lens-shaped regions, where the vertices of the nth order bound lie on the arcs of the bound of order $n-1$. When ε_1 and ε_2 are real and positive, these vertices become the upper and lower bounds on ε^*. The real and complex bounds are attained in general by anisotropic composites (MILTON, 1981a). The bounds of BAKER (1969) differ only in that we bound $E(s)$ as well as $F(s)$, so that our bounds are tighter than his. As $n \to \infty$, the bounds for fixed s converge to a point in the complex plane. Physically, this means that if one knows all the correlation functions of the material, then the effective parameter is completely characterized. Mathematically, this reflects the fact that knowledge of all the moments of a positive measure on a compact set serves to completely determine that measure.

We will now show that the above iteration procedure gives the same bounds on F as does the direct procedure discussed in GP1, which goes as follows. Knowledge of $F(s)$ to nth order,

$$F(s) = \frac{a_1}{s} + \frac{a_2}{s^2} + \dots + \frac{a_n}{s^n} + \dots, \tag{A.14}$$

is of course equivalent to knowing the first n moments of μ in (3.15) with $\mu^{(0)} = a_1, \mu^{(1)} = a_2, \dots, \mu^{(n-1)} = a_n$. Then for fixed s, $F(s,\mu)$ in (3.5) is a linear mapping from

$$M(\mu^{(0)}, \mu^{(1)}, \dots, \mu^{(n-1)}) = \left\{\mu \in M : \int_0^1 z^j \mu\,(\mathrm{d}z) = \mu^{(j)}, \quad j = 0, 1, \dots, n-1\right\} \tag{A.15}$$

to C. Now the extreme points of $M(\mu^{(0)}, \mu^{(1)}, \dots, \mu^{(n-1)})$ are the n-point measures

$$\mu\,(\mathrm{d}z) = \sum_{k=1}^{n} \alpha_k \delta_{z_k}(\mathrm{d}z), \tag{A.16}$$

where

$$\alpha_k \geqslant 0, \quad 0 \leqslant z_k < 1, \quad \sum_{k=1}^{n} \alpha_k z_k^j = \mu^{(j)}, \quad j = 0, \dots, n-1. \tag{A.17}$$

Thus extreme points of the image of $M(\mu^{(0)}, \dots, \mu^{(n-1)})$ under $F(s,\mu)$ are attained by

$$F(s) = \sum_{k=1}^{n} \frac{\alpha_k}{s - z_k} \tag{A.18}$$

where the α_k and z_k in (A.18) vary according to the moment equations in (A.17). We will show how the iteration gives the same extrema.

We first give the image of $M(\mu^{(0)}, \dots, \mu^{(n-1)})$ under the transformation

$$F_1 = \frac{1}{a_1} - \frac{1}{sF(s)}, \tag{A.19}$$

where

$$F_1 = \int_0^1 \frac{\mu_1\,(\mathrm{d}z)}{s-z}. \tag{A.20}$$

Assuming (A.14) we obtain by expansion of (A.19)

$$F_1 = \frac{b_1}{s} + \frac{b_2}{s^2} + \ldots + \frac{b_{n-1}}{s^{n-1}} + \ldots, \tag{A.21}$$

where for $1 \leqslant j \leqslant n-1$,

$$b_j = \frac{a_{j+1}}{a_1^2} - \frac{\sum_{\substack{i_1+i_2=j+2 \\ i_1,i_2 \geqslant 2}} (a_{i_1} a_{i_2})}{a_1^3} + \frac{\sum_{\substack{i_1+i_2+i_3=j+3 \\ i_1,i_2,i_3 \geqslant 2}} (a_{i_1} a_{i_2} a_{i_3})}{a_1^4} - \ldots + \frac{(-1)^{j+1} a_2^j}{a_1^{j+1}}. \tag{A.22}$$

Thus the image of $M(\mu^{(0)}, \ldots, \mu^{(n-1)})$ is $M(\mu_1^{(0)}, \ldots, \mu_1^{(n-2)})$, where $\mu_1^{(0)} = b_1$, $\mu_1^{(1)} = b_2, \ldots, \mu_1^{(n-2)} = b_{n-1}$. How the residues and poles of extreme points of $M(\mu^{(0)}, \ldots, \mu^{(n-1)})$ transform under F_1 will be made clear in what follows.

To illustrate the equivalence of the two methods, fix $s > 1$ and consider the minimization of $F(s)$ subject to

$$F(s) = \frac{a_1}{s} + \frac{a_2}{s^2} + \frac{a_3}{s^3} + \frac{a_4}{s^4} + \frac{a_5}{s^5} + \ldots. \tag{A.23}$$

The argument that we give immediately extends to complex s and general n. From the above considerations of the direct procedure, the minimum of $F(s)$ is attained within the class of measures

$$\mu\,(\mathrm{d}z) = \sum_{k=1}^{5} \alpha_k \delta_{z_k} \tag{A.24}$$

which satisfy (A.17) with $n = 5$ and $a_j = \mu^{(j-1)}$, $1 \leqslant j \leqslant 5$. Clearly

$$F(s) = \sum_{k=1}^{5} \frac{\alpha_k}{s - z_k} \tag{A.25}$$

is minimized by taking as many of the α_k and z_k equal to zero as possible. Thus the minimum is attained by

$$F(s) = \frac{\alpha_1}{s_1 - z_1} + \frac{\alpha_2}{s_2 - z_2} + \frac{\alpha_3}{s}, \tag{A.26}$$

where the α_k and z_k satisfy

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= a_1, \\ \alpha_1 z_1 + \alpha_2 z_2 &= a_2, \\ \alpha_1 z_1^2 + \alpha_2 z_2^2 &= a_3, \\ \alpha_1 z_1^3 + \alpha_2 z_2^3 &= a_4, \\ \alpha_1 z_1^4 + \alpha_2 z_2^4 &= a_5. \end{aligned} \tag{A.27}$$

Now we solve the problem via the transformation F_1. Under (A.23), F_1 has by (A.22) the expansion

$$F_1 = \frac{b_1}{s} + \frac{b_2}{s^2} + \frac{b_3}{s^3} + \frac{b_4}{s^4} + \ldots, \tag{A.28}$$

where

$$\begin{aligned} b_1 &= \frac{a_2}{a_1^2}, \quad b_2 = \left(\frac{a_3}{a_1^2} - \frac{a_2^2}{a_1^3}\right), \quad b_3 = \left(\frac{a_4}{a_1^2} - \frac{2a_2 a_3}{a_1^3} + \frac{a_2^3}{a_1^4}\right), \\ b_4 &= \left(\frac{a_5}{a_1^2} - \frac{2a_2 a_4 + a_3^2}{a_1^3} + \frac{3a_2^2 a_3}{a_1^4} - \frac{a_2^4}{a_1^5}\right). \end{aligned} \tag{A.29}$$

From the above considerations, $F_1(s)$ is minimized by

$$F_1(s) = \frac{\beta_1}{s - y_1} + \frac{\beta_2}{s - y_2}, \tag{A.30}$$

where

$$\begin{aligned} \beta_1 + \beta_2 &= b_1, \\ \beta_1 y_1 + \beta_2 y_2 &= b_2, \\ \beta_1 y_1^2 + \beta_2 y_2^2 &= b_3, \\ \beta_1 y_1^3 + \beta_2 y_2^3 &= b_4. \end{aligned} \tag{A.31}$$

We map back to F via

$$F = \frac{a_1}{s_1} \frac{1}{(1 - a_1 F_1)}, \tag{A.32}$$

which for (A.30) gives

$$F = \frac{a_1}{s} \left[1 \Big/ \left(1 - a_1 \left(\frac{\beta_1}{s - y_1} + \frac{\beta_2}{s - y_2} \right) \right) \right], \tag{A.33}$$

or

$$F = \frac{a_1 (s - y_1)(s - y_2)}{s\,[s^2 - (y_1 + y_2 + a_1(\beta_1 + \beta_2))s + (y_1 y_2 + a_1(\beta_1 y_2 + \beta_2 y_1))]}. \tag{A.34}$$

The rational function in (A.34) has a partial fraction expansion

$$F = \frac{\alpha_1}{s - z_1} + \frac{\alpha_2}{s - z_2} + \frac{\alpha_3}{s}, \tag{A.35}$$

where z_1 and z_2 are the roots of the second order polynomial in the denominator of (A.33). In order to relate the α_k and z_k from (A.35) to those from the direct method, we first expand the denominator in (A.34),

$$F = \frac{a_1}{s} \left\{ 1 \Big/ \left[1 - a_1 \left(\frac{\beta_1 + \beta_2}{s} + \frac{\beta_1 y_1 + \beta_2 y_2}{s^2} + \frac{\beta_1 y_1^2 + \beta_2 y_2^2}{s^3} + \frac{\beta_1 y_1^3 + \beta_2 y_2^3}{s^4} + \ldots \right) \right] \right\}. \tag{A.36}$$

By (A.31),

$$F = \frac{a_1}{s} \left\{ 1 \Big/ \left[1 - a_1 \left(\frac{b_1}{s} + \frac{b_2}{s^2} + \frac{b_3}{s^3} + \frac{b_4}{s^4} + \ldots \right) \right] \right\}. \tag{A.37}$$

In powers of $1/s$, (A.37) becomes

$$F = \frac{a_1}{s} + \frac{c_2}{s^2} + \frac{c_3}{s^3} + \frac{c_4}{s^4} + \frac{c_5}{s^5} + \ldots, \tag{A.38}$$

where

$$\begin{aligned} c_2 &= a_1^2 b_1, \quad c_3 = a_1^2 b_2 + a_1^3 b_1^2, \quad c_4 = a_1^2 b_3 + a_1^3 2 b_1 b_2 + a_1^4 b_1^3, \\ c_5 &= a_1^2 b_4 + a_1^3 (2 b_1 b_3 + b_2^2) + a_1^4 3 b_1^2 b_2 + a_1^5 b_1^4, \end{aligned} \tag{A.39}$$

and for general n we have when $2 \leqslant j \leqslant n$,

$$c_j = a_1^2 b_{j-1} + a_1^3 \sum_{i_1+i_2=j-1} (b_{i_1} b_{i_2}) + a_1^4 \sum_{i_1+i_2+i_3=j-1} (b_{i_1} b_{i_2} b_{i_3}) + \ldots + a_1^j b_1^{j-1}. \tag{A.40}$$

Using (A.29) in (A.39) shows that $c_2 = a_2$, $c_3 = a_3$, $c_4 = a_4$, and $c_5 = a_5$. That this must be the case can also be seen by comparing (A.38) to (A.23). That $c_j = a_j$ for general n can be shown by induction on j, i.e. by assuming $c_2 = a_2, c_3 = a_3, \ldots, c_{j-1} = a_{j-1}$ and showing $c_j = a_j$. The equations $c_2 = a_2, \ldots, c_{j-1} = a_{j-1}$ can be used in (A.40) to write the right hand side as a first order polynomial in $b_1, \ldots, b_{j-1}$ whose coefficients are powers of the a_k. Equation (A.22) can then be used to obtain $c_j = a_j$. Now we expand the partial fraction (A.35),

$$F = \frac{\alpha_1+\alpha_2+\alpha_3}{s} + \frac{\alpha_1 z_1+\alpha_2 z_2}{s^2} + \frac{\alpha_1 z_1^2+\alpha_2 z_2^2}{s^3} + \frac{\alpha_1 z_1^3+\alpha_2 z_2^3}{s^4} + \frac{\alpha_1 z_1^4+\alpha_2 z_2^4}{s^5} + \ldots \quad \text{(A.41)}$$

Comparing this with (A.38) forces the α_k and z_k to satisfy (A.27), so that (A.35) with these α_k and z_k gives the same solution as the direct method.

Since the above argument holds for general n, we have shown that the bound on F induced by F_{n-1} in (A.12) under the assumption of (A.14) is the same as that obtained from the direct method. This is so because we can successively apply the argument to F_2 with F_1 replacing F, and then to F_3 with F_2 replacing F_1, and so on. Our argument shows that at each stage the bounds obtained are optimal, i.e. they coincide with the direct bounds, so that the induced bound on F is optimal. It is interesting to note that the expressions for the α_k and z_k forced by the moment equations in (A.17) are much more complicated than the corresponding expressions found, for instance, in (A.13). Whereas the α_k and z_k are solutions of high order polynomial equations, the $\mu_i^{(0)}$ in (A.13) are ratios of polynomials in the a_k, which can be expressed as ratios of determinants of matrices containing the a_k. The complicated nature of the α_k and z_k can also be seen by viewing them as the residues and poles in the partial fraction expansion of the continued fraction (A.13). Then, for example, the z_k are the roots of high order polynomials in s. We see then that the iteration procedure and its resulting continued fraction solves the extremization problem in a convenient way.

J. Mech. Phys. Solids Vol. 34, No. 4, pp. 359–394, 1986.
Printed in Great Britain.

0022–5096/86 \$3.00+0.00
Pergamon Journals Ltd.

DISPLACIVE PHASE TRANSFORMATIONS IN SOLIDS

R. D. JAMES

Department of Aerospace Engineering and Mechanics, University of Minnesota, Mineapolis, MN 55455, U.S.A.

(*Received* 8 *July* 1985; *in revised form* 4 *November* 1985)

ABSTRACT

WE STUDY diffusionless transformations in solids which involve a sudden change of shape at a certain temperature. We assume the existence of a free energy which depends on the local change of shape and the temperature. Properties of this function reflect the underlying symmetry of the parent and product phases and an exchange of stability from parent to product phase as the body is cooled through the transformation temperature θ_0. We concentrate on two questions: (i) How can loads be applied to cause the body to transform to a particular variant of the product phase at or above θ_0? (ii) Can the parent phase be recovered by applying some system of loads at or below θ_0?

Theory and experiment are compared for thermoelastic martensitic transformations in shape-memory materials and for the α–β transformation in quartz.

1. INTRODUCTION

IN THIS paper we study the mechanical behavior of diffusionless transformations which involve a spontaneous change of shape of a crystal at a certain temperature. These transformations are termed martensitic in metals and polymorphic in other substances. Among the diffusionless transformations, "displacive" refers to transformations having a nonzero spontaneous change of shape with little hysteresis as the crystal is slowly cycled through the transformation temperature. "Reconstructive" refers to transformations having large hysteresis loops. On the molecular level, displacive transformations involve co-operative movements of atoms or groups of atoms which are not hindered by large energy barriers. Usually there is a change of symmetry in a displacive transformation; the higher symmetry usually occurs in the high temperature phase. While small hysteresis accompanies a displacive transformation, it is not necessarily true that the transformation strain or latent heat is small. Also, transformation temperatures may be altered hundreds of °C by the application of loads.

The aim of this paper is to predict the effects of load and loading device on the transformation temperature and on the stability and arrangement of the phases in some simple loading devices. In general terms, our stability criterion and constitutive equations come from Gibbs' (1875–1878) chapter on solids in contact with fluids. In this chapter Gibbs gives a finite deformation theory for the equilibrium of stressed solids in contact with fluids. He assumes that the internal energy per unit reference