Wave-field phase singularities: The sign principle

Isaac Freund and Natalya Shvartsman

Jack and Pearl Resnick Institute of Advanced Technology and Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel

(Received 23 May 1994)

Phase singularities (topological charges, dislocations, defects, vortices, etc.), which may be either positive or negative in sign, are found in many different types of wave fields. We show that on every zero crossing of the real or imaginary part of the wave field, adjacent singularities *must* be of opposite sign. We also show that this "sign principle," which is unaffected by boundaries, leads to the surprising result that for a given set of zero crossings, fixing the sign of *any given* singularity automatically fixes the signs of *all other* singularities in the wave field. We show further how the sign of the *first* singularity created during the evolution of a wave field determines the sign of *all subsequent* singularities and that this first singularity places additional constraints on the future development of the wave function. We show also that the sign principle constrains how contours of equal phase may thread through the wave field from one singularity to another. We illustrate these various principles using a computer simulation that generates a random Gaussian wave field.

PACS number(s): 42.55.−f, 42.60.Jf, 42.62.−b

I. INTRODUCTION

Phase singularities were introduced into optics (and other wave fields) by Nye and Berry in an important series of papers [1–7] that provide the theoretical foundation for the study of these fascinating, ubiquitous objects. These singularities are points from which contours of constant phase in the transverse x-y plane radiate outward in a starlike fashion such that the phase φ circulates by $2n\pi$ over any closed path that encircles the singularity. Nye and Berry [1] developed, *inter alia*, the *model* phase singularities $F_{\pm}^{(n)}=(x\pm iy)^n=r^n\exp(\pm in\theta)=A^{(n)}\exp(i\varphi_{\pm})$. The sign of the singularity is positive (F_+) if the phase $\varphi_+=+n\theta$ circulates counterclockwise and negative (F_-) if the phase $\varphi_-=-n\theta$ circulates clockwise. Since the wave function must be everywhere single valued, n is restricted to integer values. For the same reason, at the point singularity itself the amplitude $A^{(n)}=r^n$ is forced to zero.

Appending the sign (+ or −) of the singularity to n yields a quantity usually referred to as the topological charge. One reason for this designation is that in free space the phase singularities can only be created or destroyed in such a way that the total topological charge is conserved [1,3]. Generally, this means that the phase singularities are created as twins with topological charges that have opposite signs but equal magnitudes. We will find it convenient to refer to this topologically determined rule, which reflects the fact that the wave function must be everywhere continuous, as the "twin" principle. In the presence of boundaries where the wave field changes abruptly, however, the twin principle no longer necessarily holds and isolated singularities may be created or destroyed at a boundary.

Phase singularities have recently been extensively studied in nonlinear optics and laser physics, where they arise as the so-called spiral solutions to the complex Ginzburg-Landau equation [8–20], while related singularities have long been studied by condensed-matter physicists in material systems [21,22]. The phase singularities are also of the greatest importance in the linear scattering of coherent waves from random media [2,15,23–27]. In a random Gaussian speckle pattern, for example, the number density of singularities is so extraordinarily high that on average each speckle spot is accompanied by a phase singularity [2,23–26]. Important correlations between these phase singularities have been derived by Halperin in a fundamental paper that considers the statistical mechanics of topological defects [28]. The *deterministic* results we obtain here are later shown to be in full accord with Halperin's *statistical* predictions. We note that at present the tendency in optics is to use the term "vortices" to describe the phase singularities with their vortexlike phase circulation and so we adopt this term.

At some fixed value of z, the transverse component of the general wave function may be written $F(x,y)=f_{\rm Re}(x,y)+if_{\rm Im}(x,y)$. This description is particularly apt for the far field, where the internal structure of the wave field remains invariant under propagation and simply expands uniformly on the surface of a large sphere of radius z. Single valuedness of the wave function requires the centers (singular points) of any vortices that are present to lie at the intersections of the zero crossings of $f_{\rm Re}$ and $f_{\rm Im}$ [the set of points or continuous curves in which $f(x,y)=0$]. As becomes apparent, the topology of these zero crossings determines the structure of the vortices and it will prove convenient to denote a zero crossing of $f_{\rm Re}$ ($f_{\rm Im}$) by $Z_{\rm Re}$ ($Z_{\rm Im}$). Throughout this paper these zero crossings will be one of our major concerns.

Although the twin principle is of the greatest importance, when a new vortex twin is born, this principle fails to specify which twin is positive and which is negative. We have discovered a principle that fills this gap, which we call the sign principle.

1050-2947/94/50(6)/5164(9)/$06.00

Sign principle. Adjacent vortices on any given zero crossing of f_{Re} or f_{Im} must be of opposite sign.

This very general principle is unaffected by the presence of boundaries and appears to be applicable to all types of wave fields.

As a simple application of the sign principle consider a segment of Z_{Re} that contains two adjacent vortices $(A)(B)$ with signs $(+)(-)$. If a vortex twin is created between (A) and (B), the sign principle requires the configuration $(A)(-)(+)(B)$ and forbids the configuration $(A)(+)(-)(B)$. The twin principle, in contrast, allows either configuration. We note that the twin principle is a statement about the wave function at a single point (the point where vortices are created) and is thus a *local* principle, while the sign principle establishes relationships between the wave function at different points and is thus a *global* principle. Accordingly, the sign principle cannot be deduced from the twin principle. On the other hand, most applications of the twin principle do immediately follow from the sign principle.

The sign principle has many far reaching implications, some of which we develop in the following sections. In Sec. II we use simple two-color maps to describe the real and imaginary parts of the wave function, and for a given set of zero crossings, we deduce the following three propositions.

(i) The sign of any single vortex in a wave field automatically fixes the signs of all other vortices in the wave field.

(ii) If the sign of any vortex in a wave field is changed, then the signs of all other vortices must also be changed.

(iii) The sign of the first vortex created during the evolution of a wave field will fix the signs of all future generations of vortices.

We call these three propositions the three corollaries of the sign principle.

Following the elucidation of the three corollaries, we give in Sec. III a proof of the sign principle itself. We start by considering in Sec. III A a wave field that contains only first-order zeros. We relax this restriction in Sec. III B and show how to apply the sign principle to wave fields with higher-order zeros. We also show in this section how the sign principle may be used to analyze ambiguous situations in order to determine the topological charge of unusual singularities. In Sec. III C we introduce an important extension of the sign principle that permits its application to wave fields containing isolated zero crossings and we show how the future evolution of a wave field will be strongly constrained by the very first vortices that are created. In Sec. IV we show that the sign principle also strongly constrains how the equiphases in a wave field may thread from one vortex to another and we present three "phase rules" that summarize these constraints. In Sec. V we briefly discuss possible extensions of the research and we compare the (deterministic) sign principle with Halperin's calculations of the statistical correlations of vortices [28]. Throughout, we illustrate and test the various propositions using the results of a computer simulation that generates a random Gaussian wave field [26,27].

II. THE THREE COROLLARIES OF THE SIGN PRINCIPLE

As we are interested in physically real wave functions, we assume that both f_{Re} and f_{Im} are regular, single-valued functions of the spatial coordinates x and y. Fractals and other interesting monstrosities are therefore excluded. Returning to the basic description of a (linear) vortex $F_{\pm}=x\pm iy$, we imagine standing on the zero crossing of f_{Re} (the y axis) while orienting ourselves such that the $+x$ axis is to our right. If we always adopt this convention, then f_{Re} always increases to our right and decreases to our left. For the positive vortex F_{+}, when we stand on the zero crossing of f_{Im}, the x axis, f_{Im} increases to our front and decreases to our back, while for the negative vortex F_{-}, f_{Im} decreases to our front and increases to our back. Thus the sign of a vortex is determined by the directions of increase (or decrease) of f_{Re} and f_{Im} and this remains true also for the general wave function.

As an aid in determining in which direction f_{Re} or f_{Im} increases or decreases in the general case, we prepare a two-dimensional, two-color map for each of these functions. In coloring these maps we adopt the convention that the positive regions of a function are colored white and the negative regions are colored black. Examples of maps of f_{Re} and f_{Im} for a random wave field obtained from our computer simulation [26,27] are shown in Fig. 1. The boundaries between the white and black regions on these maps are the zero crossings Z_{Re} and Z_{Im}. As before, we imagine standing at the origin of a local coordinate system with $+x_{\mathrm{local}}$ to our right and $+y_{\mathrm{local}}$ to our front. By our convention, when we stand on Z_{Re} we always orient ourselves such that white is to our right, so that $\partial f_{\mathrm{Re}}/\partial x_{\mathrm{local}}>0$ and $\partial f_{\mathrm{Re}}/\partial y_{\mathrm{local}}=0$. The sign of a vortex is determined by the (coordinate-independent) *sign* of the Jacobian $\partial(f_{\mathrm{Re}},f_{\mathrm{Im}})/\partial(x_{\mathrm{local}},\partial y_{\mathrm{local}})$ [26,28]. Thus, within our convention the sign of a vortex is the sign of $\partial f_{\mathrm{Im}}/\partial y_{\mathrm{local}}$.

FIG. 1. Maps of (a) f_{Re} and (b) f_{Im} for a random Guassian wave field. Positive regions of the functions are colored white and negative regions are colored black. The zero crossings of these maps and their associated vortices are shown in Fig. 6.

In the highly unlikely event that one or both of our maps contain isolated points at which two apparently different white regions just contact one another, ambiguities in how we are to orient ourselves, and thus ambiguities in the values of the local derivatives, may arise. These ambiguities may be resolved by introducing a small, continuous, *local* perturbation of the wave function that slightly raises or lowers the function at the contact point. Such a perturbation preserves, of course, the topology of the wave field. If we raise the wave function, then we join the two regions together to produce a single white region, while if we lower the wave function, we separate the two regions. Having eliminated all ambiguities in this way we complete our analysis and at the end relax the perturbation back to zero. Continuity, together with the fact that our perturbation of the wave function is completely local, guarantees the same end result no matter which route we choose for eliminating the contact point. Thus, for a given topology of $Z_{\rm Re}$ and $Z_{\rm Im}$ the signs of the vortices are completely determined by the relative colorings of our maps.

We now show that the converse of the above is also true and that given $Z_{\rm Re}$ and $Z_{\rm Im}$, the sign of *any* single vortex automatically determines the relative colorings of all regions on both maps and thus the signs of *all other* vortices. We begin by noting the obvious point that our map of $f_{\rm Re}$, for example, *must* change color every time we cross a $Z_{\rm Re}$ and *must not* change color if we do not cross a $Z_{\rm Re}$. Accordingly, given the color of *any small* finite region, the colors of *all other* regions in the map are uniquely determined by the requirement that the wave function be everywhere single valued (uniquely colored). Maps for which contradictory colorings arise via different paths correspond to wave functions that are not single valued and need therefore be discarded as unphysical. All of the above is also obviously true for the map of $f_{\rm Im}$.

We will now assume here and throughout the remainder of this paper that the *only* information about the wave field available to us *a priori* is a zero crossing map showing all the $Z_{\rm Re}$ and $Z_{\rm Im}$. At a given intersection of $Z_{\rm Re}$ and $Z_{\rm Im}$ on our zero crossing map, if we know the sign of the vortex that is present at this intersection, we know the relative local colorings of the (unavailable) maps of $f_{\rm Re}$ and $f_{\rm Im}$. For example, if we *arbitrarily* decide that one side of $Z_{\rm Re}$ on the $f_{\rm Re}$ map is colored white (black), then the sign of our vortex determines which side of $Z_{\rm Im}$ on the $f_{\rm Im}$ map must also be colored white (black). But as just discussed, once we have fixed the coloring of a small finite region on the map of $f_{\rm Re}$ and on the map of $f_{\rm Im}$, we have fixed the colorings everywhere and thus the signs of all vortices in the wave field. Accordingly, we obtain the first corollary of the sign principle: (i) The sign of any single vortex in a given zero crossing map uniquely determines the signs of all other vortices. Further, if we switch the color of a region containing a vortex on only *one* of our maps, thereby switching the sign of that vortex, the requirement that the wave function be everywhere single valued forces us to switch the colors of all other regions of that same map and this in turn changes the signs of all other vortices. Accordingly, we have the second corollary of the sign principle: (ii) Changing the sign of any single vortex in a wave field automatically changes the signs of all other vortices in the wave field. Finally, if we add new zero crossings to our maps, then the colors of all new regions created are already predetermined, as are the signs of all new vortices. We thus obtain the third important corollary of the sign principle: (iii) The sign of the first vortex created will determine the signs of all subsequent vortices that may be created.

III. THE SIGN PRINCIPLE

A. First-order zeros

We turn now to a proof of the sign principle. For simplicity, we start with the assumption that both $Z_{\rm Re}$ and $Z_{\rm Im}$ do not contain any self-intersections. Formally, this is equivalent to assuming that both $f_{\rm Re}$ and $f_{\rm Im}$ are "suitably" regular at level zero [29], so that both $Z_{\rm Re}$ and $Z_{\rm Im}$ are everywhere smooth, continuous, and differentiable. Our assumption requires that all vortices be first-order zeros since, as we show in Sec. III B where we relax this assumption, higher-order zeros necessarily give rise to self-intersections.

We first note that as we move along any continuous segment of $Z_{\rm Re}$ we never encounter a point at which we suddenly need to turn around in order to keep white to our right. It was to ensure this result that we first eliminated all self-intersections. Making our way along $Z_{\rm Re}$, we carefully record the value of $f_{\rm Im}$ as a function of our position along the contour. We denote the resulting function by $f_{\rm Im}(l_{\rm Re})$ in order to emphasize that it is a *one*-dimensional function of the arc length $l_{\rm Re}$ along $Z_{\rm Re}$. Plotting this function we observe (Fig. 2) that its slope $df_{\rm Im}/dl_{\rm Re} = \partial f_{\rm Im}/\partial y_{\rm local}$ always alternates in sign from

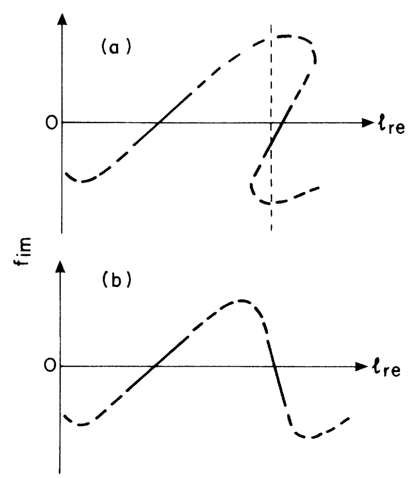

FIG. 2. $f_{\rm Im}$ as a function of arc length $l_{\rm Re}$ along $Z_{\rm Re}$. (a) Adjacent zero crossings of $f_{\rm Im}$ have slopes with the same sign, causing the wave function to be multivalued (vertical dashed line). (b) Adjacent zero crossings of $f_{\rm Im}$ have slopes with opposite signs, as required for a single-valued function.

one of its zero crossings to the next. This is a direct consequence of the requirement that $f_{\rm Im}$ must be everywhere single valued. Since the sign of a vortex is determined by the sign of $\partial f_{\rm Im}/\partial y_{\rm local}$, we have established that adjacent vortices on any (and therefore every) $Z_{\rm Re}$ *must* alternate in sign. Repeating the whole chain of argument given above while everywhere replacing "right" with "left" and $f_{\rm Re}$ with $f_{\rm Im}$ establishes that also adjacent vortices on every $Z_{\rm Im}$ *must* alternate in sign. We thus arrive at the general rule, which we call the sign principle: Adjacent vortices on any zero crossing must alternate in sign. Starting at any vortex whose sign is known, this rule lets us rapidly skim along $Z_{\rm Re}$ or $Z_{\rm Im}$ (or when convenient cross back and forth between contours) labeling vortices alternately plus or minus as we go. In Sec. III B we show how to apply the sign principle to wave fields containing higher-order zeros and in Sec. III C we extend the sign principle to the case of (usually closed) $Z_{\rm Re}$ and $Z_{\rm Im}$ contours that are isolated from all other contours on our maps. With these extensions, the sign principle then provides a simple algorithm for the practical implementation of corollaries (i) and (iii).

We note that the above arguments are unaffected by boundaries that may abruptly terminate a zero crossing, so that unlike the twin principle, the sign principle holds also in the presence of boundaries. The reason for this important difference between the two principles is that boundaries may make the wave function discontinuous, but never multivalued.

B. Higher-order zeros

If our wave field contains higher-order zeros, say, $F_+^{(2)}=[(x-x_0)+i(y-y_0)]^2=|\mathbf{r}-\mathbf{r}_0|^2\exp(+2i\theta)$, which corresponds to a doubly degenerate vortex with topological charge $+2$, then we have self-intersections for both $Z_{\rm Re}$ and $Z_{\rm Im}$. Here $\theta=\arctan[(y-y_0)/(x-x_0)]$, and both $f_{\rm Re}=(x-x_0)^2-(y-y_0)^2$ and $f_{\rm Im}=2(x-x_0)(y-y_0)$ correspond to saddle points that just touch the x-y plane at the vortex center $x_0,y_0(\mathbf{r}_0)$. Since self-intersections imply that two white regions on our map touch at a single point (a contact point), they must be eliminated. As before, we accomplish this by adding a local perturbation to the wave function, replacing $f(x,y)$, for example, by $f(x,y)+\varepsilon\exp\{-[(x-x_0)^2+(y-y_0)^2]/w^2\}$. We may turn the perturbation continuously on and off by varying ε, while by adjusting w we can make the perturbation as local as we wish. As shown in Fig. 3, such a small perturbation splits a self-intersecting contour into two self-avoiding contours, the details of which depend upon the sign of ε. For either choice of sign, since we have eliminated the self-intersections we may freely apply the sign principle to the perturbed wave field. Having done this we collapse ε back to zero with the assurance that by continuity and the locality of the perturbation, our results for vortex signs cannot depend upon whether we initially chose ε to be positive or negative. Zeros of arbitrary (finite) higher order may all be easily handled in this same fashion. But as is obvious from the example in Fig. 3, the degeneracy of a vortex is ultimately not important, and in applying

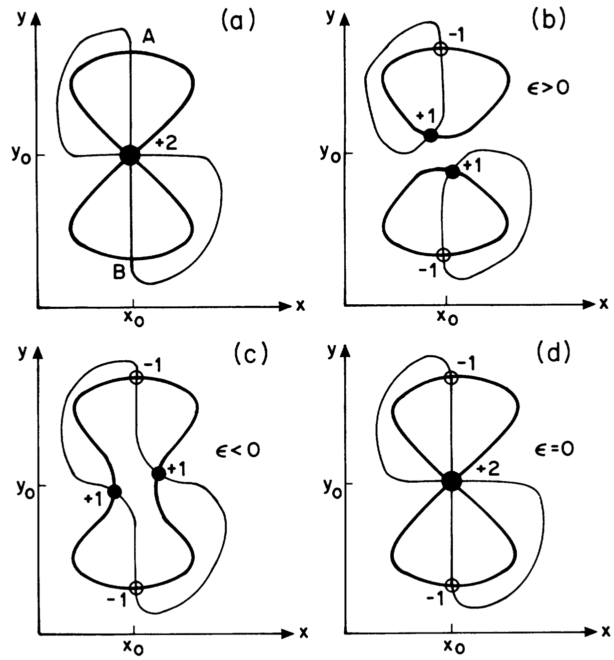

FIG. 3. Application of the sign principle to a higher-order zero. (a) Zero crossings $Z_{\rm Re}$ (——) and $Z_{\rm Im}$ (——) of a doubly degenerate vortex with topological charge $+2$ centered at x_0,y_0. Both $Z_{\rm Re}$ and $Z_{\rm Im}$ have a self-intersection at the vortex center that introduces apparent ambiguities in the determination of the signs of vortices A and B using the sign principle. (b) and (c) The self-intersections are turned into avoided crossings by the introduction of a small local perturbation ε. (b) $\varepsilon>0$. (c) $\varepsilon<0$. For both choices of ε the degeneracy is lifted, the central vortex splits into two separated positive vortices (●) with topological charges $+1$, and the sign principle determines that vortices A and B are both negative (○). (d) The wave field after the perturbation ε is relaxed back to zero. This illustrates that the sign principle may be applied to degenerate vortices using only their *sign* without regard to the magnitude of their topological charge.

the sign principle to degenerate vortices we need not bother to actually split self-intersections and may instead simply use the *sign* of the vortex without regard to its degeneracy.

The sign principle may also prove useful in resolving ambiguous situations. Suppose we have $f_{\rm Re}=x^2-y^2$ and $f_{\rm Im}=y$. Does this correspond to a vortex entered on the origin, and if so, what is the vortex sign and what is its topological charge? As shown in Fig. 4, we once again turn the self-intersection of $Z_{\rm Re}$ into an avoided crossing. If we do this by making $\varepsilon>0$, $Z_{\rm Re}$ and $Z_{\rm Im}$ do not intersect at all, and so there is certainly no vortex at the origin. On the other hand, if we make $\varepsilon<0$ we end up with a closely spaced positive-negative vortex pair (topological charge ±1) symmetrically disposed about the origin. In either case we may now freely apply the sign principle to the perturbed wave field. When we collapse ε back down to zero the vortex pair created for $\varepsilon<0$ self-anihilates, so that for either choice of sign for ε we conclude that our original wave function did not contain a vortex at the origin, but rather some other sort of phase discontinuity.

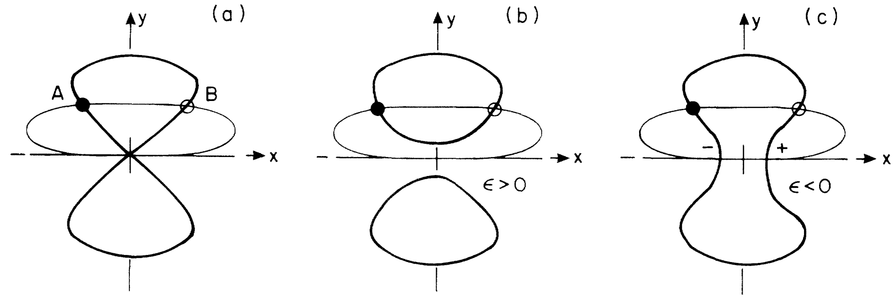

FIG. 4. Resolution of an ambiguous phase discontinuity using the sign principle. (a) A self-intersection of Z_{Re} (——) also intersects Z_{Im} (——) at the origin. From the form of the wave function, vortex A is positive (●) and vortex B is negative (○). Since there is no possible sign for the phase discontinuity at the origin that can satisfy the sign principle, this discontinuity cannot correspond to a vortex. (b) and (c) A small local perturbation ε applied to f_{Re} turns the self-intersection of Z_{Re} into an avoided crossing, thereby facilitating application of the sign principle. (b) $\varepsilon > 0$. Z_{Re} and Z_{Im} do not intersect, confirming that there is no vortex at the origin. (c) $\varepsilon < 0$. A positive-negative vortex pair is created, which self-annihilates when ε is relaxed back to zero, again confirming that the original wave function does not contain a vortex at the origin.

C. The extended sign principle

We now extend the sign principle to include also the case of noncontacting, apparently isolated zero crossings, an example of which is shown in Fig. 5. In Fig. 5(a) we show a single isolated positive vortex, labeled A. As already discussed, because of the presence of boundaries the twin principle cannot be invoked and the creation of a single vortex is permitted [14]. In Fig. 5(b) we add a second set of intersecting zero crossings and create a second vortex, labeled B. Although corollaries (i) and (iii) ensure us that the sign of vortex B is predetermined, since there is no common zero crossing contour that connects vortex A to vortex B, the sign principle in its present form cannot be used to find the sign of vortex B. We could, of course, work out the relative colorings of maps of f_{Re} and f_{Im} in order to find the sign of vortex B, but there is an easier method. We *extend* the sign principle by introducing a continuous distortion of any one of the contours containing vortex A, say, Z_{Re}, until we cross a Z_{Im} contour containing vortex B. Having done this we apply the sign principle in the usual way, finding thereby that the sign of vortex B is positive. This is illustrated in Fig. 5(c). After fixing the sign of vortex B we undo the distortion of Z_{Re} and return to our original zero crossing map. We emphasize that it is quite irrelevant whether or not the wave field actually evolves to the configuration shown in Fig. 5(c). Since this configuration maintains the topology of Z_{Re} and is generated *continuously* from the initial state by a completely *local* perturbation, it and the initial state must have the same vortex structure (map colorings). Of course, any other contour distortion that meets these requirements will also yield the same final result. We illustrate this procedure further in Fig. 6, which is taken from our computer simulations for a random Gaussian wave field [26,27]. We note that since our simulation provides us with complete maps of Z_{Re} and Z_{Im} together with the signs of all vortices, we have been able to directly verify the sign principle in hundreds of specific cases, including dozens of instances that require the extended principle involving contour distortions. Worth emphasizing is that in performing contour distortions it is best to avoid introducing intersections of Z_{Re} (Z_{Im}) contours with other Z_{Re} (Z_{Im}) contours (i.e., self-

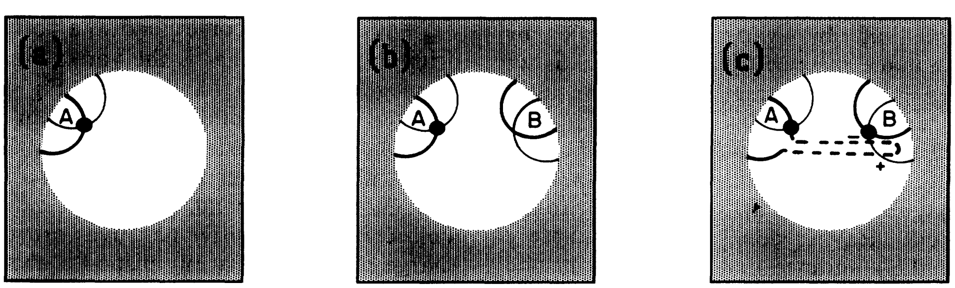

FIG. 5. Extending the sign principle. (a) A single positive (●) vortex A is created at a boundary [Z_{Re} (——), Z_{Im} (——)]. (b) A second vortex B appears whose zero crossing contours do not overlap the contours of vortex A so that the sign principle cannot be applied directly. (c) Distortion (— — —) of the Z_{Re} contour of vortex A that permits the sign principle to be applied to vortex B. The signs of the virtual vortices (+ or −) created by this distortion follow immediately from the sign principle and the sign of vortex B is now easily seen to be positive.

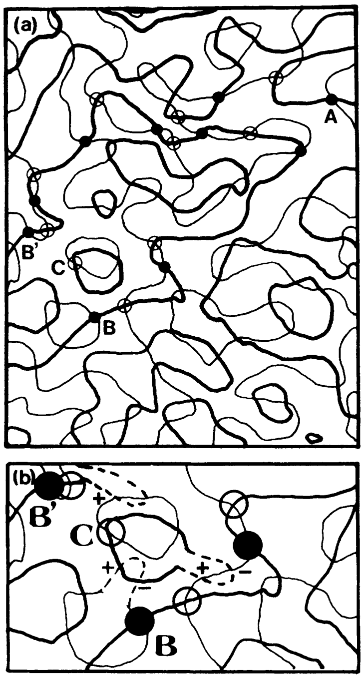

FIG. 6. (a) Zero crossing map of the random Gaussian wave field shown in Fig. 1 [Z_{Re} (——),Z_{Im} (——)]. Starting at positive vortex A, the sign principle is used to determine vortex signs along the two paths A-B and A-B', with ● denoting positive vortices and ○ negative vortices. (b) The sign of vortex C is determined to be negative using the extended sign principle (— — —). The virtual vortices created by the contour extensions also obey the sign principle and are labeled + or −. As required, all three contour extensions yields the same sign for vortex C.

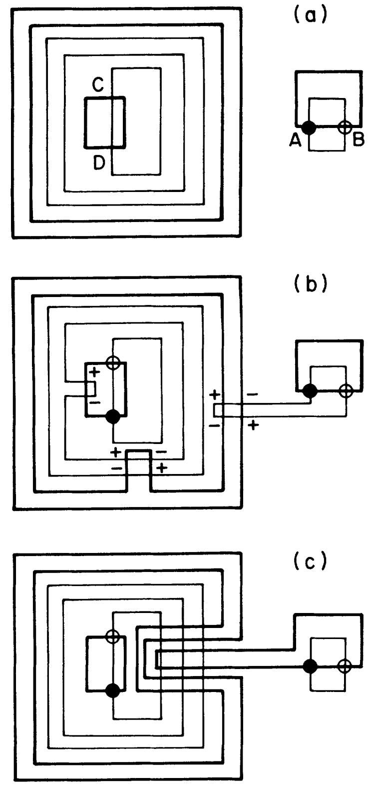

FIG. 7. Application of the extended sign principle to isolated, enclosed vortices [Z_{Re} (——),Z_{Im} (——)]. (a) Vortex A is positive (●) and vortex B is negative (○). The signs of vortices C and D are to be determined using the extended sign principle. (b) and (c) Examples of contour extensions that permit the sign of vortex C to be determined as negative and the sign of vortex D as positive. The virtual vortices created by the contour extensions in (b) are labeled + or −. The reader may enjoy labeling the virtual vortices created in (c).

intersections). In Fig. 7 we provide examples that illustrate ways of doing this that should suffice to handle any zero crossing map we may encounter. Finally, we note that the extended sign principle is truly *global* in scope, since starting at any given point in the wave field it can reach out to *every* other point in the wave field.

We now illustrate how the first vortex created during the evolution of a wave field already, via the single principle, strongly constrains the possible future evolution of the wave field. In Fig. 8(a) we show a zero crossing map containing two vortices A and B. Fixing *a priori* the sign of vortex A as positive, the extended sign principle fixes also vortex B as positive. We now introduce another set of intersecting zero crossings to this map, as in Fig. 8(b), and we ask what is the sign of vortex C, the new vortex created at the intersection? In order to answer this question we again apply the extended sign principle. Starting from the contours that house vortex A we conclude that vortex C must be negative, but using the contours that house vortex B leads us to the opposite conclusion that vortex C must be positive. Since there is nothing in Fig. 8 that voids the sign principle, we are forced to conclude that the new contours introduced in Fig. 8(b) are impossible, i.e., the wave field *cannot* evolve from Fig. 8(a) to Fig. 8(b). In light of the random, seemingly arbitrary zero crossings in Fig. 6, from which we might be tempted to infer that "anything goes," the conclusion that Fig. 8(b) does not "go" is rather unexpected. The correctness of this conclusion is confirmed by preparing colored maps of f_{Re} and f_{Im} for Fig. 8(b) and observing that if vortices A and B have the same sign, then there is no possible coloring of these maps that does not require contradictory colors for some region. Accordingly, Fig. 8(b) corresponds to a wave function that is not single valued and is forbidden. [In Figs. 8(c) and 8(d) we use the sign principle to show that it is the new Z_{Re} contour which is "illegal."] As the number of vortices increases, the number of constraints on possible forms of the wave function grows rapidly, so we may conclude that to a large extent the vortices determine much of the structure of a wave field. This result is in accord with our previous conclusions based upon very different arguments that employed the sampling theorem [25].

IV. PHASES: RULES AND CONJECTURES

Since the phase of the wave is $\varphi = \arctan(f_{\mathrm{Im}}/f_{\mathrm{Re}})$, the zero crossing contour Z_{Im} is also a contour of constant

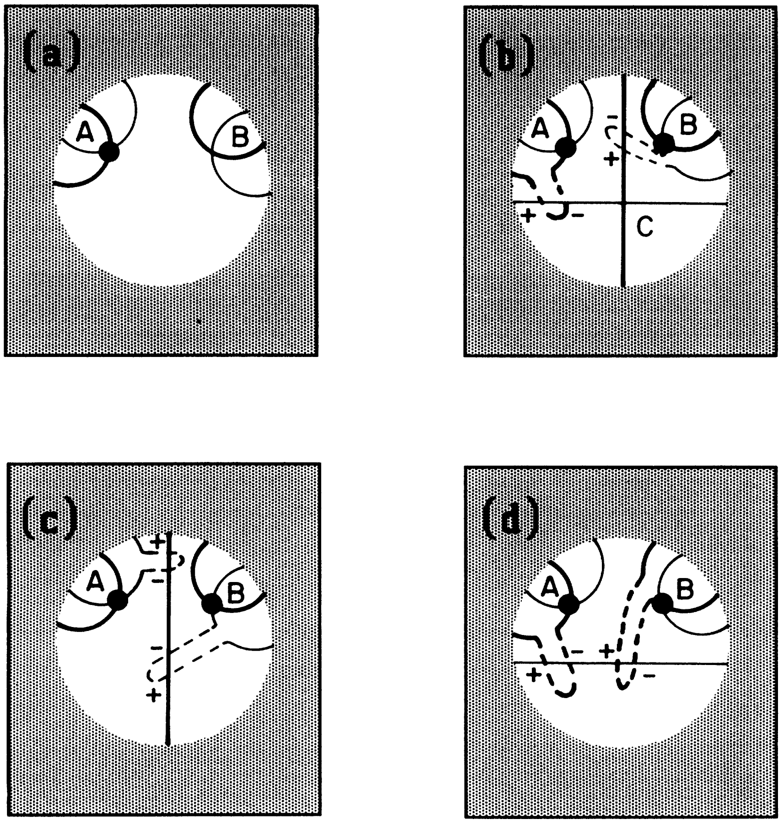

FIG. 8. Evolution of a wave field is constrained by its vortices [Z_{Re} (——), Z_{Im} (——)]. (a) Positive (●) vortex A is created first. The extended sign principle fixes as also positive the sign of vortex B, which is created next. (b) New Z_{Re} and Z_{Im} evolve creating vortex C. Starting from vortex A the extended sign principles fixes the sign of vortex C as positive, while starting from vortex B the sign of vortex C is found to be negative. This contradiction implies that at least one of the newly evolved zero crossings corresponds to a multivalued wave function and is "illegal." (c) and (d) The newly introduced zero crossings are tested for legality using contour distortions (— — —). (c) Test of the new Z_{Re}. Since the virtual vortices that appear violate the sign principle (two adjacent vortices are negative) the new Z_{Re} is illegal and must be discarded. (d) Test of the new Z_{Im}. The virtual vortices that appear obey the sign principle. The new Z_{Im} is legal and may be retained.

phase $\varphi=0$ or π and Z_{Re} is a contour of constant phase $\varphi=\pi/2$ or $3\pi/2$. As already discussed, we may always locally distort the wave field in order to temporarily eliminate zeros of order higher than one, and since for vortices which are first-order zeros the phase always changes discontinuously by π on any line passing through the vortex center, it will be convenient to sometimes use reduced phases $\varphi^*=\varphi(\mathrm{mod}\ \pi)$. We thus have that Z_{Im} is a reduced equiphase with $\varphi^*=0$ and Z_{Re} is a reduced equiphase with $\varphi^*=\pi/2$.

Now, the phase of an optical wave may be easily shifted by passing the wave through a uniform thickness glass plate. Equivalent uniform phase shifters exist for other wave fields. Accordingly, *any* equiphase in a wave field may always be turned into a zero crossing. But simply shifting the phase of the wave does not change the zeros of amplitude, so the positions of the vortex centers remain unchanged. As mentioned, the sign of a vortex may be conveniently calculated from the sign of the Jacobian $\partial(f_{\mathrm{Re}},f_{\mathrm{Im}})/\partial(x,y)$ evaluated at the vortex center [26,28,30]. Direct calculation verifies that this Jacobian is, as expected, also invariant to a uniform phase shift, so we are led to the first phase rule.

(i) Adjacent vortices on any (and therefore every) reduced equiphase in a wave field must of opposite sign.

Since every vortex has a unique sign, we obtain the second phase rule.

(ii) No equiphase can begin and end on the same vortex.

From (i) and (ii) follows the third phase rule.

(iii) All equiphases that begin at a vortex must end on vortices of opposite sign or continue on to the boundaries of the wave field.

Just like the sign principle from which they derive, these phase rules may be directly applied to degenerate vortices (higher-order zeros) using only the *sign* of the vortex without regard to the magnitude of its topological charge. It is thus apparent that the vortices are very much analogous to electric charges and the equiphases to electric-field lines. This is in full accord with our previous results for model multivortex wave functions [25]. The phase rules, however, are very general and are independent of analogies or models. We emphasize that these rules do not prohibit closed equiphases elsewhere in the wave field.

If there are only two vortices present and these have opposite signs, then all the equiphases of the first vortex that do not reach the wave-field boundary must end on the second vortex. When there are many vortices with different signs, however, different equiphases that begin on a given vortex can, and often will, end on different vortices. We will refer to two vortices that are threaded by one or more reduced equiphases as being "connected." A vortex that is connected to more than one neighbor will be described as "multiconnected" and sets of vortices in which each member of the set is connected to all other members of the set will be referred to as "interconnected." After examining how the equiphases thread from vortex to vortex in a random Gaussian wave field we are led to the following two conjectures: (i) the *weak* phase conjecture and (ii) the *strong* phase conjecture.

(i) All vortices in a random Gaussian wave field are multiconnected with probability approaching one.

(ii) All vortices in a random Gaussian wave field are interconnected with probability approaching one.

The strong phase conjecture implies that starting at *any* given vortex in the wave field, *any other* vortex in the wave field may always be reached by moving along some reduced equiphase.

We note that unlike the sign principle, which refers to what happens along a line (zero crossing) and whose proof was therefore reducible to a simple statement about the behavior of regular functions in *one* dimension, the proofs of these phase conjectures represent truly *two*-dimensional problems. Unfortunately, as emphasized by Adler [29], problems in the geometry (topology) of random fields that cannot be reduced to one dimension are rarely solved.

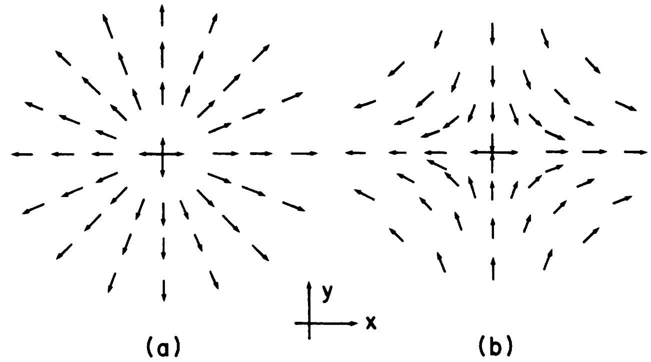

FIG. 9. Vector field ("spin") representation of the order parameter. The x component of a (unit) vector equals $f_{\rm Re}/(f_{\rm Re}^2+f_{\rm Im}^2)^{1/2}$, the y component equals $f_{\rm Im}/(f_{\rm Re}^2+f_{\rm Im}^2)^{1/2}$, and the angle that the vector makes with the x axis is the value of the order parameter (phase) at the point x,y. (a) Positive vortex $x+iy$. The vectors "flow" radially outward (positively) from the vortex center. (b) Negative vortex $x-iy$. The vectors flow inward (negatively) along the y axis and outward (positively) along the x axis. The negative vortex is thus not a simple "sink" for the positive vortex "source."

V. DISCUSSION

Vortices are normally described as being found at the intersections of the zero crossings of the real and imaginary parts of the field. We have seen, however, that these zero crossings are arbitrary and may be replaced by any pair of equiphases that differ in phase by $\pi/2$. But even this latter restriction is unnecessary, since at the intersection of *any* two equiphases the wave function becomes multivalued unless the amplitude goes to zero. Accordingly, the most general statement about vortices is that they are located at the intersections of *any* two equiphases.

Although our major interest has been (polarized) random optical waves, the results described here are relevant to all regular (scalar) wave fields, such as sound waves, electron waves, neutron waves, etc. Our results should also prove applicable to any continuous, two-dimensional, two-component order parameter that contains topological singularities. This wide class of order parameters describes many different thin-film *material* systems [21,22,28]. Here the zero crossings are either those of the real and imaginary parts of the order parameter in a complex representation or of its x and y components in a vector representation. In optics one describes the vortices in terms of their equiphases, but in condensed-matter physics one conventionally uses a field of unit vectors ("spins") in which the orientations of the vectors describe the spatial dependence of the order parameter (phase). In Fig. 9 we display this representation. The vector field for the simple positive vortex $x+iy$ may be seen to be intuitively satisfying, but this is no longer the case for the simple negative vortex $x-iy$ [21]. In multivortex wave fields more complex vortex morphologies such as $F(x,y)=(a_{\rm Re}+b_{\rm Re}x+c_{\rm Re}y)+i(a_{\rm Im}+b_{\rm Im}x+c_{\rm Im}y)$ are the rule rather than the exception ($-\infty\le a,b,c\le\infty$) [25,26] and the corresponding rather complicated vector fields become even less intuitive and harder to interpret. Accordingly, it appears that the vector field representation of the order parameter may not provide a useful alternative for deriving the sign principle and the other general propositions developed here. Nonetheless, since the vector field is widely used in condensed-matter physics, we generalize our results also to this representation. Defining a reduced order parameter (equivalent to the reduced phase φ^*) by removing the heads of the arrows in Fig. 9, we have, for any two-dimensional system described by a continuous two-component order parameter, the following proposition.

Proposition. Adjacent vortices on any (and therefore every) contour of constant reduced order parameter must alternate in sign.

We may speculate that results that are analogous to this proposition, which includes both the sign principle and the phase rules, may exist also in dimensions that are higher than two and for order parameters that have more than two components, and we would suggest that a search for such extensions may prove rewarding.

When applicable, the sign principle implies that the contribution of vortex signs to the system's entropy is negligible ($k_B\ln 2$, where k_B is Boltzmann's constant, instead of $N_v k_B \ln 2$, where N_v is the number of vortices), so that vortex statistics, thermodynamics, phase transitions, etc. may all be strongly modified by this new principle. We note in this regard that Halperin, in his important paper on the statistical mechanics of topological defects [28] to which we referred earlier, has presented a fundamental calculation of the ensemble average topological charge correlation function for a random, two-dimensional Gaussian wave field. Halperin finds that positive vortices tend to be surrounded by negative vortices and vice versa. This finding is, of course, in full accord with the sign principle, which provides a *deterministic* basis for these *statistical* correlations. We note that our measurements of the vortex topological charge correlation function, which will be reported on separately [31], are in substantial agreement with Halperin's calculations, thereby confirming not only the applicability of these calculations, but also the importance of the sign principle in determining the statistical mechanics of vortices.

For two-dimensional wave fields, the sign principle and its corollaries, together with the phase rules and their conjectures, imply that the vortices form a set that in a real sense is highly ordered. This set, which extends throughout the whole wave field, clearly determines much if not all of the wave-field structure [8,10,14,16,26]. The existence of so much hidden order in a random Gaussian wave field, for example, is quite surprising, since one normally thinks of this wave field as being composed of *independent* coherence areas (regions of nearly constant amplitude and phase). This point of view is clearly untenable as far as the vortices are concerned. In a separate paper we describe our studies on Gaussian speckle patterns that confirm some dozen different correlations between vortices predicted by the sign principle [31]. Since the vortices and their interconnected equiphases determine both the locations and the sizes of the regions of slowly varying phase normally associated with speckle spots, the independence of these spots is now also called into question. It appears likely that our current picture of random wave fields will need to be reexamined.

In their pioneering work, Nye and Berry suggested that the arrangement of vortices in a wave field could serve as a useful means for classifying nontrivial wave-field structures [1–7]. In solid-state physics, for example, many properties of a crystal are already revealed simply by stating the crystal class. In order to actually carry out Nye and Berry's proposal for wave-field classification, however, one needs to know all the different possible interconnections between vortices. The results given here may provide a useful start in this direction.

We end with an unanswered question. In multiple-scattering media when the photon-scattering mean free path becomes of order the wavelength of light, many strong correlations of the intensity in widely different regions of the wave field set in and ultimately the wave becomes localized [32–35]. What happens to the vortices?

ACKNOWLEDGMENT

We are pleased to acknowledge the support of The Israel Science Foundation of the Israel Academy of Sciences and Humanities.

[1] J. F. Nye and M. V. Berry, Proc. R. Soc. London Ser. A **336**, 165 (1974).
[2] M. V. Berry, J. Phys. A **11**, 27 (1978).
[3] F. J. Wright, in *Structural Stability in Physics,* edited by W. Guttinger and H. Eikemeier (Springer-Verlag, Berlin, 1979), p. 141.
[4] M. Berry, in *Physics of Defects,* edited by R. Balian, M. Kleman, and Jean-Paul Poirier (North-Holland, Amsterdam, 1981), p. 453.
[5] J. F. Nye, Proc. R. Soc. London Ser. A **378**, 219 (1981).
[6] F. J. Wright and J. F. Nye, Philos. Trans. R. Soc. London Sect. A **305**, 339 (1982).
[7] *Polarization* singularities have been discussed by J. F. Nye, in *Physics of Defects* (Ref. [4]), p. 545; Proc. R. Soc. London Ser. A **387**, 105 (1983); **389**, 279 (1983); J. F. Nye and J V. Hajnal, *ibid.* **409**, 21 (1987); J. V. Hajnal, *ibid.* **414**, 433 (1987); **414**, 447 (1987).
[8] P. Coullet, L. Gil, and J. Lega, Phys. Rev. Lett. **62**, 1619 (1987).
[9] P. Coullet, L. Gil, and F. Rocca, Opt. Commun. **73**, 403 (1989).
[10] G. Goren, I. Procaccia, S. Rasenat, and V. Steinberg, Phys. Rev. Lett. **63**, 1237 (1989).
[11] F. T. Arecchi, G. Giacomelli, P. L. Ramazza, and S. Residori, Phys. Rev. Lett. **65**, 2531 (1990).
[12] L. Gil, J. Lega, and J. L. Meunier, Phys. Rev. A **41**, 1138 (1990).
[13] M. Brambilla, F. Battipede, L. A. Lugiato, V. Penna, F. Prati, C. Tamm, and C. O. Weiss, Phys. Rev. A **43**, 5090 (1991); **43**, 5114 (1991).
[14] F. T. Arecchi, G. Giacomelli, P. L. Ramazza, and S. Residori, Phys. Rev. Lett. **67**, 3749 (1991).
[15] P. L. Ramaza, S. Residori, G. Giacomelli, and F. T. Arecchi, Europhys. Lett. **19**, 475 (1992).
[16] G. Indebetouw and S. R. Liu, Opt. Commun. **91**, 321 (1992).
[17] G. S. McDonald, K. S. Syed, and W. J. Firth, Opt. Commun. **94**, 469 (1992).
[18] G. A.Swartzlander, Jr. and C. T. Law, Phys. Rev. Lett. **69**, 2503 (1992).
[19] R. Neubecker, M. Kreuzer, and T. Tschudi, Opt. Commun. **96**, 117 (1993).
[20] L. Gil, Phys. Rev. Lett. **70**, 162 (1993).
[21] J. M. Kosterlitz and D. J. Thouless, in *Progress in Low Temperature Physics,* edited by B. D. F. Brewer (North-Holland, Amsterdam, 1978), Vol VII, p. 371.
[22] N. D. Mermin, Rev. Mod. Phys. **51**, 591 (1979).
[23] N. B. Baranova, B. Ya. Zel'dovich, A. V. Mamaev, N. Pilipetskii, and V. V. Shkukov, Pis'ma Zh. Eksp. Teor. Fiz. **33**, 206 (1981) [JETP Lett. **33**, 195 (1981)].
[24] N. B. Baranova, A. V. Mamaev, N. Pilipetskii, V. V. Shkunov, and B. Ya. Zel'dovich, J. Opt. Soc. Am. **73**, 525 (1983).
[25] I. Freund, N. Shvartsman, and V. Freilikher, Opt. Commun. **101**, 247 (1993).
[26] I. Freund, J. Opt. Soc. Am. A **11**, 1644 (1994).
[27] N. Shvartsman and I. Freund, Phys. Rev. Lett. **72**, 1008 (1994).
[28] B. I. Halperin, in Physics of Defects (Ref. [4]), p. 814.
[29] R. J. Adler, *The Geometry of Random Fields* (Wiley, New York, 1981), Chap. 3.
[30] G. E. Volovik and V. P. Mineev, Zh. Eksp. Teor. Fiz. **72**, 2256 (1977) [Sov. Phys. JETP **45**, 1186 (1977)].
[31] N. Shvartsman and I. Freund, J. Opt. Soc. Am. A **11**, 2710 (1994).
[32] S. John, Phys. Rev. Lett. **53**, 2169 (1984).
[33] P. W. Anderson, Philos. Mag. B **52**, 505 (1985).
[34] A. Z. Genack and N. Garcia, Phys. Rev. Lett. **66**, 2064 (1991).
[35] S. John, Phys. Today **44**(5), 32 (1991).

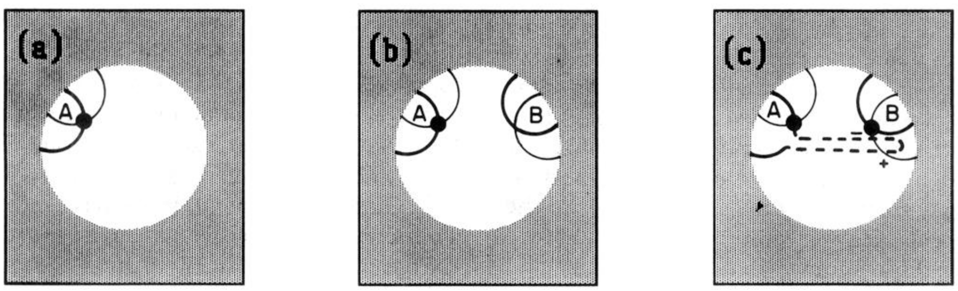

FIG. 5. Extending the sign principle. (a) A single positive (●) vortex A is created at a boundary [Z_{Re} (——), Z_{Im} (——)]. (b) A second vortex B appears whose zero crossing contours do not overlap the contours of vortex A so that the sign principle cannot be applied directly. (c) Distortion (— — —) of the Z_{Re} contour of vortex A that permits the sign principle to be applied to vortex B. The signs of the virtual vortices (+ or −) created by this distortion follow immediately from the sign principle and the sign of vortex B is now easily seen to be positive.

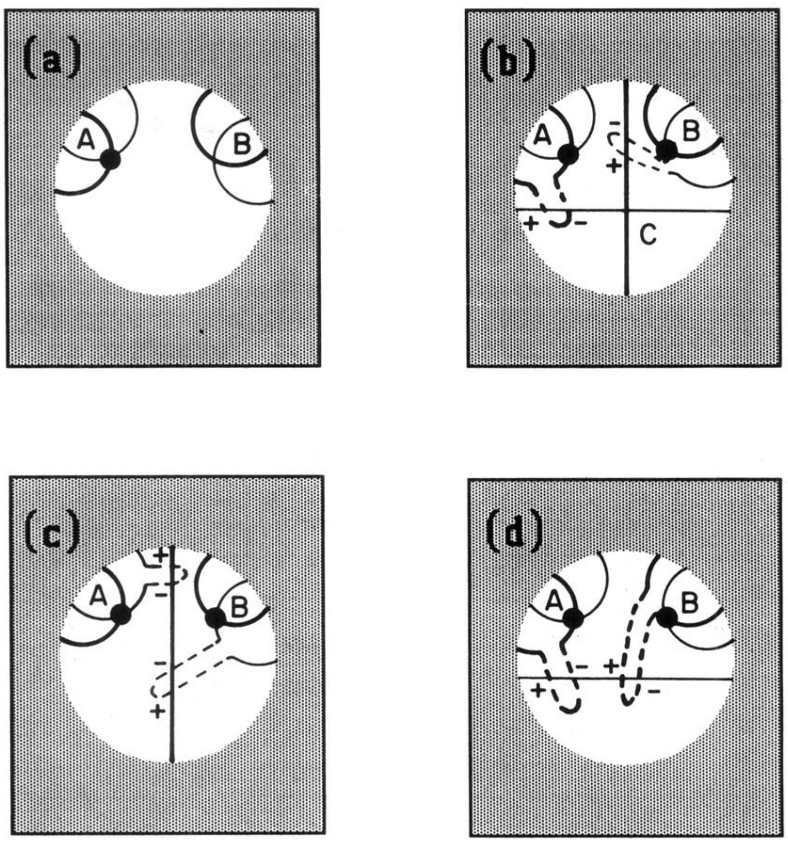

FIG. 8. Evolution of a wave field is constrained by its vortices [Z_{Re} (——),Z_{Im} (——)]. (a) Positive (●) vortex *A* is created first. The extended sign principle fixes as also positive the sign of vortex *B*, which is created next. (b) New Z_{Re} and Z_{Im} evolve creating vortex *C*. Starting from vortex *A* the extended sign principles fixes the sign of vortex *C* as positive, while starting from vortex *B* the sign of vortex *C* is found to be negative. This contradiction implies that at least one of the newly evolved zero crossings corresponds to a multivalued wave function and is "illegal." (c) and (d) The newly introduced zero crossings are tested for legality using contour distortions (— — —). (c) Test of the new Z_{Re}. Since the virtual vortices that appear violate the sign principle (two adjacent vortices are negative) the new Z_{Re} is illegal and must be discarded. (d) Test of the new Z_{Im}. The virtual vortices that appear obey the sign principle. The new Z_{Im} is legal and may be retained.