Theory of the vortex breakdown phenomenon

By T. BROOKE BENJAMIN

Department of Engineering, University of Cambridge†

(Received 11 June 1962)

The phenomenon examined is the abrupt structural change which can occur at some station along the axis of a swirling flow, notably the leading-edge vortex formed above a delta wing at incidence. Contrary to previous attempts at an explanation, the belief demonstrated herein is that vortex breakdown is not a manifestation of instability or of any other effect indicated by study of infinitesimal disturbances alone. It is instead a finite transition between two dynamically conjugate states of axisymmetric flow, analogous to the hydraulic jump in open-channel flow. A set of properties essential to such a transition, corresponding to a set shown to provide a complete explanation for the hydraulic jump, is demonstrated with wide generality for axisymmetric vortex flows; and the interpretation covers both the case of mild transitions, where an undular structure is developed without the need arising for significant energy dissipation, and the case of strong ones where a region of vigorous turbulence is generated. An important part of the theory depends on the calculus of variations; and the comprehensiveness with which certain properties of conjugate flow pairs are demonstrable by this analytical means suggests that present ideas may be useful in various other problems.

1. Introduction

The term 'vortex breakdown', or the alternative 'vortex bursting', is commonly used to refer to the abrupt and drastic change of structure which sometimes occurs in a swirling flow, particularly in the leading-edge vortex formed above a sweptback lifting surface. The first experimental observations were made independently at about the same time four years ago by several people investigating the aerodynamics of delta wings, notably by W. E. Gray, R. L. Maltby, N. C. Lambourne and T. Elle, and accounts have been published by Werlé (1960) and Elle (1960). Frequent reference will be made presently to the experimental paper by Harvey (1962) which appears in this issue of the *Journal of Fluid Mechanics*; and for information on the practical problem in whose context wide interest in the vortex breakdown phenomenon has arisen, one may usefully refer to the analysis of the leading-edge vortex published recently by Hall (1961).

There have already been several attempts to account for the phenomenon theoretically. Jones (1960) and Ludwieg (1961) proposed it to be the outcome

† This paper was written while the author was on leave at the Institute of Science and Technology, University of Michigan.

of instability of the original flow and they examined the stability of relevant flow models by means of small-disturbance theory on the usual lines. While it is not to be denied that instability in the sense commonly understood may be responsible for the disruption of many swirling flows in practice, the view to be emphasized in this paper is that the distinctive vortex breakdown phenomenon is not such a case and its essential explanation is outside the reach of conventional stability theory. One outstanding piece of evidence controverting the previous explanation is that under careful experimental conditions the phenomenon can be made approximately axisymmetric (Harvey 1962), while the original flow is of a kind that is highly stable to axisymmetric disturbances (cf. Jones 1960). The additional facts that the breakdown can then comprise an abrupt expansion of the stream surfaces near the axis, and moreover be an approximately steady configuration, are also incompatible with the previous explanation.

Approaching the problem quite differently Squire (1960) suggested that breakdown might occur when the flow can sustain infinitesimal standing waves, his idea being that, if such waves exist, disturbances which are present far downstream might spread along the vortex and hence disrupt the flow nearer the start. Since standing waves of indefinitely great length are the first to become possible as the velocity of swirl is gradually increased, he proposed the limiting condition for the existence of such waves to be the inceptive state for a vortex breakdown. Experimentally the breakdown phenomenon is indeed observed to move up from downstream on inception. But a serious objection to Squire's theory is that the group velocity of his standing waves is in fact directed downstream (see §3 below), which means that the waves can only form in the rear of a disturbing agency and cannot spread upstream (i.e. they are like gravity waves in a horizontal water channel). However, the analytical condition for the existence of indefinitely long standing waves has a fundamental role in the present theory, even though it does not refer directly to any physical event. A hint as to the proper interpretation of Squire's analysis, together with a correct conjecture about the real nature of vortex breakdown, has in fact already been given by Harvey (1960).

The present explanation for vortex breakdown recognizes it as an example of a general type of fluid-dynamical mechanism which apparently has several other instances with practical significance,† and whose essentials can be described neatly and comprehensively in the language of the calculus of variations. It will be shown that frictionless swirling flows generally occur in conjugate pairs, both of which can form parts of the same overall system. For a given distribution of total head and circulation over the stream surfaces, one possible state of flow is 'subcritical' in the sense that infinitesimal axisymmetric standing waves can occur upon it, and the conjugate state is 'supercritical'. (The case considered by Squire is, of course, the 'critical' one separating these two domains of flow states.) A deduction of paramount importance in the theory is that, compared with their conjugates, supercritical flows possess a deficiency of total momentum,

† For example, the application of present ideas to horizontal shear flows of heterogeneous fluids will be obvious, and there appear to be possible applications to flows in magnetic fields.

or more precisely of total thrust or 'flow-force' defined as the integral of axial momentum flux plus pressure over a section through the flow. For reasons to be argued later, this property implies that supercritical flows are liable to undergo spontaneous transitions to the subcritical state; the corresponding gain of flow-force is either manifested as the 'wave resistance' of a stationary wave-train formed in the subcritical flow downstream from the transition, or it induces such a violent wave-making action that turbulence results and this dissipative process cancels the original dynamical relationships determining the subcritical conjugate. The end result of either mechanism is that the flow-force downstream is reduced to the upstream value, a steady state then prevailing in conformity with the momentum balance between sections either side of the transition. The latter mechanism is necessary when the original flow is far supercritical, so that the transition requires a drastic change of conditions and the excess flow-force associated with the 'intact' conjugate state is large; but wave formation may balance most of the excess without significant energy loss when the transition is a mild one. (Note the evidence of a periodic structure in the photographs presented by Harvey (1962), which show a mild case of vortex breakdown.)

In all these respects vortex breakdown is precisely analogous to the hydraulic jump in open-channel flow; and because of its capacity to illustrate the essentials of the explanation for vortex breakdown, the familiar hydraulic problem is reviewed in §2. While the circumstantial distinctions between undular and fully turbulent hydraulic jumps are well-recognized experimental facts, however, the proper theoretical interpretation is far from obvious and does not appear to be widely known, so that there is a need to discuss the matter rather carefully in order to secure the basic ideas which carry over intact to vortex breakdown.

In §3 some relevant aspects of the linearized theory of perturbed cylindrical vortices are covered, and in particular analytical criteria for supercritical and subcritical states of flow are derived—which later, in §4, are identified with conditions for the flow-force integral to be a minimum or not. This section serves mainly to establish a number of preliminary details needed for the main analysis in §4, and it may safely be skipped in a first reading of the paper—except perhaps for the short subsection 3.3. Where basic properties of axisymmetric swirling flows are concerned, the derivations of several results presented in §§3 and 4 are deferred to the Appendix at the end of the paper. Although the displacement of a fairly large amount of this supporting material to the Appendix makes it an unusually long one, this form of presentation seemed specially desirable to give clarity to the essential points of the theory, which need not be regarded as tied exclusively to the particular context of vortex flows.

2. The hydraulic analogy

The simple theory of the hydraulic jump, which is due to Rayleigh (see Lamb 1932, p. 280), considers a steady transition between two states of uniform perfect-fluid flow along a horizontal open channel. Presuming conservation of flow rate and momentum (corrected for pressure force), it leads to the conclusion that energy must be dissipated at the transition. The matter can be argued in a rather different way, however, which indeed is necessary to explain the fact that very weak hydraulic jumps comprise stationary wave-trains showing no evidence of significant energy loss (see Brooke Benjamin & Lighthill 1954); and this approach will be outlined here on account of its close affinity with our explanation for vortex breakdown.

We first recall that for a given flow rate Q and total head H there are two possible states of uniform steady flow along a channel of rectangular section, each having constant depth h and horizontal velocity u. One is subcritical, which means that its Froude number $F = u/(gh)^{\frac{1}{2}}$ is less than unity, and the other is supercritical with $F > 1$. For the subcritical flow, h is larger and hence u is smaller. A property of all subcritical flows is that infinitesimal standing waves can occur upon them, whereas no such wave is possible for supercritical flows. These facts are indicated by the well-known formula (Lamb 1932, p. 376, equation (5))

$$F^2 = \frac{\lambda}{2\pi h}\tanh\frac{2\pi h}{\lambda} \tag{2.1}$$

for the wavelength λ of stationary sinusoidal waves, which gives real values of λ only if $F < 1$, with $\lambda \to \infty$ for $F \to 1^-$.† As mentioned already in § 1, these waves have a 'relative' group velocity (i.e. relative to phase, which is of course at rest for standing waves) which is directed downstream; consequently a wave-train will form in the rear of an obstacle disturbing a subcritical stream, but one cannot form ahead of it (cf. Lamb, § 248).

Consider next the quantity S defined as the sum of horizontal momentum flux and pressure force per unit span, thus being given by $S = \rho(u^2h + \frac{1}{2}gh^2)$ for a uniform stream with density ρ. (As frequent reference is made to this quantity throughout this paper, we introduce for it the befitting term 'flow-force', already used in § 1.) For any pair of conjugate flows as specified above, S is easily shown to be greater for the subcritical than for the supercritical one. Using suffices 1 and 2 to refer to subcritical and supercritical respectively, one gets

$$S_1 - S_2 = \frac{\rho g(h_1 - h_2)^3}{2(h_1 + h_2)}, \tag{2.2}$$

which shows the excess flow-force of the subcritical conjugate to be of third-order smallness when the difference between the pair is small. It can also be shown that

† We note a generalization of this result which will be seen to correspond to the outcome of the analysis in § 3. The wave problem providing (2.1) is concerned with infinitesimal disturbances whose dependence on the horizontal co-ordinate x is representable by $\exp(i\alpha x)$, with $\alpha = 2\pi/\lambda$ real. Consider now the more general problem of disturbances dependent on $\exp(\gamma x)$, with γ unrestricted. Formal solution of the same dynamical equations and boundary conditions leads to the result $F^2 = (\tan\gamma h)/\gamma h$. This equation has infinitely many pairs of real roots $\pm\gamma$ but, as already implied, just one pair of purely imaginary roots $\gamma = \pm i\alpha$ if $F^2 < 1$ and none if $F^2 > 1$; furthermore, it has no complex roots (Lamb, § 245). The existence of an infinite set of positive real eigenvalues γ^2, with a limited number of negative ones only under subcritical conditions, is a universal feature of the general class of systems to which the main ideas of this paper will apply. Incidentally, a significance of the least real value γ^2 is that the respective exponential functions of x will describe the outlying parts of any disturbance in a supercritical stream: for instance, the outskirts of a solitary wave or the converging stream issuing from under a sluice-gate (Brooke Benjamin 1956). A corresponding interpretation for swirling flows will be noted in § 3.

S is reduced when a train of standing waves is formed on a subcritical flow, this reduction being equivalent to the 'wave resistance' as usually defined, and S decreases steadily with increasing wave amplitude up into the realm of 'waves of finite size' and ultimately to breaking waves (cf. Brooke Benjamin & Lighthill 1954). Thus, for example, we have an explanation for each of the situations illustrated in figure 1. Here an obstacle spanning a subcritical stream exerts a horizontal force against it which must be balanced by a reduction in flow-force downstream. In figure 1 (*a*) the reduction is shown to occur by the formation of waves, and in figure 1 (*b*) a larger reduction is shown to be brought about by a transition to the conjugate supercritical state.

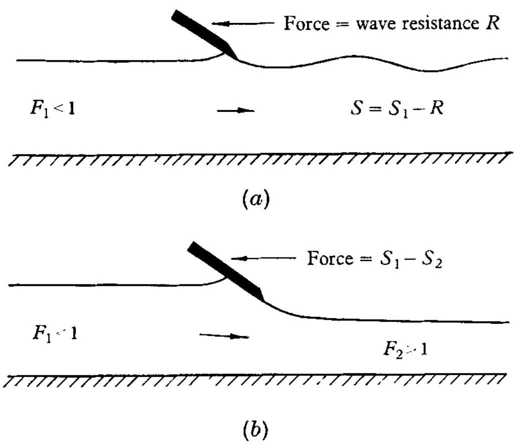

FIGURE 1. Physical illustrations of (*a*) wave resistance and (*b*) flow-force deficiency of supercritical flow.

For an originally supercritical flow, on the other hand, a transition to the conjugate subcritical state would result in an *increase* of flow-force; that is, an external force acting in the downstream direction would be necessary to push back against the rise in water level so as to restore the flow-force balance essential to a steady state. This imagined release of flow-force accounts vividly for the tendency of supercritical streams to form hydraulic jumps, and it is perhaps more appealing intuitively than Rayleigh's theoretical model for the final steady state of a jump. If a transition arises in the absence of external forces and at first there is no energy loss, the unbalance of flow-force implies, of course, that the situation cannot be steady. In fact, for a mild transition with the upstream Froude number not much greater than unity, waves develop one by one behind the transition so that eventually a steady wave-train is established on the subcritical flow, the reduction in S by wave resistance being enough to make up the momentum balance between the two flows without much energy being lost (see figure 2(*a*)).† For strong transition, however, the large excess of flow-force

† It was shown by Brooke Benjamin & Lighthill (1954) that some energy loss is in fact essential, since the only wave possible when Q, H and S are all conserved is the solitary wave; however, an exceedingly small reduction in H is enough to give the right conditions for a steady 'cnoidal' wave-train such as is observed behind a very weak hydraulic jump.

causes a violently unsteady motion leading to 'breaking' of the water surface (i.e. this is nature's unsuccessful attempt in the direction of cancelling the excess by wave formation, the initial wave being swelled to breaking still without making an adequate flow-force reduction); thus the flow becomes very turbulent at the transition, and the large energy loss thereby occurring nullifies the conditions of the present theory. To account for the final steady state in this case, one must adopt the alternative hypothesis that only Q and S are conserved,

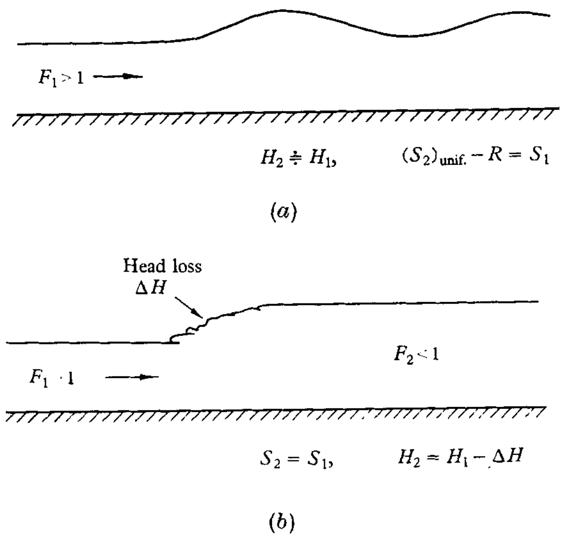

FIGURE 2. Illustration of two types of hydraulic jump: (*a*) undular, and (*b*) dissipative with uniform flow downstream.

allowing that H may be substantially decreased through the transition (see figure 2 (*b*)). Intermediate between the extreme cases of a weak undular hydraulic jump with extremely slight energy loss and a strong fully turbulent one, there is the case of a jump led by a breaking wave with smooth ones behind; a large fraction of the energy dissipation according to Rayleigh's theory may then occur, yet the effect of wave resistance may still be significant.

We shall see later that every one of the aspects of open-channel flows noted above has a counterpart in axisymmetric swirling flows, although the analytical problem for the latter is much more intricate. As soon as the corresponding set of basic properties of conjugate states has been established in general for swirling flows, the explanation for vortex breakdown will be evident as an exact parallel to the foregoing explanation for the hydraulic jump. Indeed there will hardly be need to set out the argument again in the new context.

There is another property of the hydraulic jump which evidently carries over to vortex breakdown, but which is not represented explicitly by the present theory. A hydraulic jump will arise *spontaneously* on a supercritical stream in a long uniform channel only when F is fairly close to unity (values near 2 are typical). On a highly supercritical stream, a jump will not occur unless precipitated by some special means, such as an obstruction spanning the bed of the channel, whose removal would result in the remains of the jump being swept away downstream. In other words, a jump is possible only when the situation downstream is capable of bearing the subcritical state of flow that would be developed in the rear of the jump; or, in the language of practical hydraulics, we say that the jump must be matched to the 'backwater' conditions. By analogy, the most dangerous condition with regard to the possibility of spontaneous vortex breakdown is a marginally supercritical state of flow. Nevertheless, very violent breakdowns from a far supercritical state may occur when there is some special formative agency, such as an adverse pressure gradient along the axis of the vortex.

3. Properties of perturbed cylindrical vortices

In preparation for § 4, we note some details from the linearized theory of small disturbances in frictionless axisymmetric swirling flows. Let x and r denote axial and radial co-ordinates, and let $y = \frac{1}{2}r^2$. In the primary flow, which is taken to be cylindrical and to fill the space bounded by $y = a$, the axial velocity $W(y)$ and swirl velocity $V(y)$ are prescribed functions of y alone. Accordingly, a stream-function $\Psi(y)$ can be defined such that $W = d\Psi/dy$; and the circulation

$$2\pi r V = 2\pi K(y),$$

say, and hence the quantity $I = \frac{1}{2}K^2$, are also known functions of y.

3.1. *Deductions by Sturm–Liouville theory*

Suppose now that this flow is given a stationary axisymmetric perturbation, such that the total stream-function takes the general form

$$\psi(x,y) = \Psi(y) + \epsilon\phi(y)\, e^{\gamma x}.$$

(For example, if γ is a pure imaginary, say $i\alpha$, this describes a standing wave of wavelength $2\pi/\alpha$.) Directly from the equations of motion linearized in ϵ, or alternatively as is shown in the Appendix, it is found that $\phi(y)$ must be a solution of the second-order differential equation (cf. Appendix, equation (A 19))

$$\phi_{yy} + \left\{\frac{\gamma^2}{2y} - \frac{W_{yy}}{W} + \frac{I_y}{2y^2W^2}\right\}\phi = 0; \tag{3.1}$$

and the kinematical boundary conditions on the flow (Appendix, equations (A 20)) require that

$$\phi(0) = 0, \quad \phi(a) = 0. \tag{3.2}$$

These equations comprise a Sturm–Liouville system, and the properties we first show are deduced directly from the general theory of such systems (Ince 1926, Ch. 10). The conditions imposed in the usual development of the theory are satisfied in general here, in particular since the coefficient of ϕ in (3.1) is non-singular for $0 < y \leqslant a$. [Even in the exceptional case where W has a zero in this interval, as will be considered briefly in § 4.7, the alternative form of the coefficient indicated by equation (A 14) in the Appendix shows it to be still non-singular if, as may be assumed for any realistic representation of the flow, the circulation and total-head distributions are wholly continuous; thus W_{yy} and $(I_y)^{\frac{1}{2}}$ must become proportional to W in the neighbourhood of its zero. However, in important respects to be explained in §4.7, this case is inconsistent with the collective interpretation of the properties of swirling flows which is demonstrated in the following pages, and which provides the basic explanation of vortex breakdown; accordingly, a reservation should be made about it from the start.] The coefficient behaves like y^{-1} as $y \to 0$ since necessarily $I \to 2\omega^2 y^2$, where ω is the angular velocity at the centre; but solutions of (3.1) with singular behaviour at $y = 0$ are ruled out from present consideration by the first boundary condition.

As the factor multiplying γ^2 in (3.1) is positive, the theory proves the existence of an infinite set of real eigenvalues $\gamma_0^2, \gamma_1^2, \gamma_2^2, \ldots$ (denoted here in their order along the axis of real numbers) which allow the corresponding solutions $\phi_0, \phi_1, \phi_2, \ldots$ of (3.1) to satisfy (3.2), and whose only limit-point is $\gamma^2 = +\infty$ (Ince, §10.61). Thus we see that the perturbation can always take an infinite number of forms with real exponential dependence on x (i.e. with $\gamma^2 > 0$), but there is a limited possibility of standing waves (i.e. with $\gamma^2 < 0$, so that $\gamma = i\alpha$). A given state of flow may be termed *supercritical* if all the eigenvalues γ^2 are positive, so that no standing wave is possible, and *subcritical* if at least the first one γ_0^2 is negative. Again, the state of flow investigated by Squire (1960) is *critical* in the present sense, since the condition $\gamma_0^2 = 0$ assumed by him obviously divides those just defined.

[The following physical interpretation of γ_0 in the supercritical case is worth noting incidentally. Consider a supercritical flow which is undisturbed at infinity upstream, and which is disturbed locally either by some axisymmetrical obstacle or by a breakdown to the subcritical state. Then clearly the disturbance will diminish in the upstream direction ultimately as $\exp(-|\gamma_0 x|)$, since all other possible components of a disturbance which has become small will diminish more rapidly than this (i.e. $\gamma_0 < \gamma_1$, etc.). The extremity of the 'nose' which leads a mild vortex breakdown (see the photographs presented by Harvey (1962)) might well fit this description. Note that the nose will become longer for a weaker breakdown, i.e. for upstream conditions approaching nearer the critical so that γ_0^2 becomes smaller. Conversely, a strong breakdown from a state of flow far beyond critical can only occur very abruptly.]

A further distinction between supercritical and subcritical states will now be noted, which, though primarily analytical in character, will appear later to have important physical implications. For the moment we reconsider (3.1) as if γ^2 were a free parameter and, for the particular case $\gamma^2 = 0$, take ϕ_c to denote a solution which satisfies *one* of the boundary conditions (3.2). That is, we have

$$(\phi_c)_{yy} - \left\{\frac{W_{yy}}{W} - \frac{I_y}{2y^2 W^2}\right\}\phi_c = 0 \tag{3.3}$$

throughout the interval $(0, a)$ and $\phi_c = 0$ at one of the end-points. Since (3.3) is a linear equation, ϕ_c is therefore determined completely except for an arbitrary constant multiplier. Now Sturm's fundamental comparison theorem (Ince, §10.3) shows on comparing the coefficients of (3.1) and (3.3) that, if any eigenvalue γ_n^2 of the complete system is negative, then ϕ_c oscillates more rapidly in $(0, a)$ than the respective eigenfunction ϕ_n; on the other hand, ϕ_c oscillates less rapidly than ϕ_n if γ_n^2 is positive. It follows at once that ϕ_c has at least one zero for

$0 < y < a$ if the flow is subcritical ($\gamma_0^2 < 0$), whereas ϕ_c can have no zero in this interval if the flow is supercritical. In other words, a necessary and sufficient condition for the existence of standing waves of finite length (i.e. for a subcritical state) is that the 'test function' ϕ_c should vanish at least once for $0 < y < a$.

3.2. *Generalized specification of supercritical and subcritical conditions*

At this point it seems most desirable to define a characteristic dimensionless parameter which, like the Froude number for idealized open-channel flows, will be greater or smaller than unity accordingly as the flow is supercritical or subcritical. Indeed, recognizing the generality of the basic motion of supercritical and subcritical states, one sees the need for a generalized definition independent of the details of the physical situation. We recall that the Froude number is defined in effect as the ratio of the (uniform) stream velocity to the velocity of extremely long waves relative to the fluid. For the purpose in question, however, such a definition is adequate *only* if the flow is uniform, since otherwise the absolute wave velocity—which has to vanish for long waves propagating against a critical flow—is generally not the sum of the mean flow velocity and the wave velocity found for propagation over stationary fluid (e.g. see Burns 1953). We therefore propose the following definition which will serve for every situation.

Let c_+ and c_- denote the absolute velocities, measured positively in the direction of flow, at which waves of extreme length propagate respectively with and against the flow. The general problem of finding these velocities is left to the Appendix, §d; but we may note here the fairly obvious facts that whereas c_+ is necessarily positive, c_- may have either sign, being positive when the convective action of the flow on the waves outweighs the relative propagation (the supercritical case) and negative when the waves can make headway upstream against the flow (the subcritical case). Accordingly we define

$$N = \frac{c_+ + c_-}{c_+ - c_-} \tag{3.4}$$

as our universal characteristic parameter, having that $N > 1$ specifies supercritical conditions in the general sense and $N < 1$ subcritical conditions. [In the case of the open-channel flow, we have $c_+ = (gh)^{\frac{1}{2}} + u$ and $c_- = -(gh)^{\frac{1}{2}} + u$, so that (3.4) recovers the usual definition of the Froude number $F = u/(gh)^{\frac{1}{2}}$.]

3.3. *The general effect of swirl*

Supposing that the original flow before a breakdown is always unidirectional ($W > 0$), we have that *in the absence of swirl* it would necessarily be supercritical; for with $I_y = 0$ everywhere the solution of (3.3) vanishing at $y = 0$ is

$$\phi_c = AW \int_0^a \frac{dy}{W^2},$$

which can have no zero for $y > 0$ except where $W = 0$. Moreover, when the circulation steadily increases outwards ($I_y > 0$), as is usual in the flows preceding vortex breakdowns (see Hall 1961), the influence of the swirl on the flow condition as represented by N is always in the direction from supercritical towards subcritical. This is evident from the fact that any enlargement of I_y throughout $(0, a)$ makes the coefficient in (3.3) more positive, and so makes the solution vary more rapidly—i.e. tend towards achieving a complete oscillation within $(0, a)$. By consideration of travelling waves, as in the Appendix, §*d*, it readily appears that a reduction of the swirl uniformly to zero corresponds to the limit $N \to \infty$.

Therefore, since breakdowns in practice appear to be the ultimate outcome of the amount of swirl getting larger (or rather the relative amount as measurable by the helix angle—see Harvey (1962)), we must regard the approach to breakdown conditions as, in effect, a reduction of N from high supercritical values towards the critical. Hence we base our investigation on the assumptions that the original flow is supercritical, and that breakdowns occur when N is brought down locally to some value not much greater than unity.†

Suppose, for example, that $W = \chi W(y)$ and $V = \omega V(y)$, where χ and ω are numerical factors which are varied parametrically while $W(y)$ and $V(y)$ remain the same. Dimensional reasoning shows that N will depend only on the ratio ω/χ, i.e. on the helix angle at some representative y; and it is evident from what has been written above that N will vary in some way inversely with ω/χ, so that $N \to \infty$ for $\omega/\chi \to 0$ (cf. the results for the two examples in §5, particularly equations (5.11) and (5.28)).

Some further remarks on the present aspect are made at the end of §4.6.

3.4. *Wave resistance and group velocity*

Two further results needed as details of the complete explanation for vortex breakdown have to be noted here, although the proofs of them are deferred to the Appendix. First, we consider the 'wave resistance' or flow-force reduction associated with the standing wave possible in the subcritical case, it being supposed that the process of wave formation does not alter the total head (or energy per unit mass) on any stream surface. For infinitesimal standing-wave disturbances in the form $\phi(y) \sin(\alpha x + \nu)$, equation (A 25) in the Appendix shows that the change in flow-force from the value for the undisturbed flow is independent of x, as it obviously must be on physical grounds, and is necessarily negative. This property of flow-force reduction is the counterpart of a property of gravity waves mentioned in §2.‡ Although the property is definitely established only for infinitesimal waves, it is entirely reasonable to suppose that the flow-force reduction will increase steadily with increasing wave amplitude up to finite magnitudes, just as in the known case of gravity waves. There is as yet no available theory of finite-amplitude waves in swirling flows with general velocity distributions.

The second result is that the group velocity of the standing waves is directed downstream, which is another property in common with gravity waves. There

† One cannot in general be more specific than this about the location of vortex breakdowns, just as one cannot for hydraulic jumps—see the last paragraph of §1, also the second of §6.

‡ It should not be thought that this property is common to all standing-wave systems. A contrary example is provided by capillary waves on a liquid stream; since they form in front of a fixed obstacle, the corresponding wave resistance experienced by the obstacle must be equivalent to a flow-force *increase* upstream.

is a fairly obvious physical argument suggesting that this property must be concomitant with the first; but to prove it definitely one has to consider waves travelling at velocity c and examine the limit of their relative group velocity $\alpha(dc/d\alpha)$ as $c \to 0$. This is done in the Appendix, §*d*. Subsequent to the introduction of conditions productive of waves, the wave-train will spread away from its point of origin—i.e. develop wave by wave—at a rate equal to the group velocity. This general result applies precisely only to infinitesimal waves, of course, but finite-amplitude waves can reasonably be expected to have the same direction of formation, as is known to be the case for gravity waves.

4. Theory of finite transitions between frictionless cylindrical flows

4.1. *Basic ideas*

We may regard as the fundamental dynamical properties of a steady axisymmetric flow that the total head H and circulation $2\pi K$ have fixed values on any stream-surface $\psi = \text{const.}$ (cf. Appendix, §*a*). Thus we may write

$$yV^2 = 2K^2(\psi) = I(\psi), \tag{4.1}$$

and, confining attention to situations in which there is no radial component of velocity,

$$p/\rho + \tfrac{1}{2}(W^2 + V^2) = H(\psi), \tag{4.2}$$

where p is pressure and ρ density. For a cylindrical flow as assumed, V, W and ψ are functions of y alone. The idea implied here is that (4.1) and (4.2) represent completely the dynamical conditions for the existence of a state of flow with cylindrical stream-surfaces in a given frictionless system; and obviously there can be a multiplicity of such states within the same system, each corresponding to different kinematical conditions. Moreover, we shall presently demonstrate the important fact that, even with fixed kinematical conditions, the specification of $I(\psi)$ and $H(\psi)$ does not necessarily determine a flow uniquely.

The general dynamical problem represented by (4.1) and (4.2) is now reapproached in a somewhat unusual way. Consider the following integral expressing the flow-force S of a cylindrical flow bounded by $y = a$:

$$S = 2\pi \int_0^a (\rho W^2 + p)\, dy. \tag{4.3}$$

Substituting for p from (4.2), then for V^2 from (4.1), and finally putting

$$W = \psi_y \equiv d\psi/dy,$$

we get

$$S = 2\pi\rho \int_0^a \left\{\tfrac{1}{2}\psi_y^2 + H(\psi) - \frac{I(\psi)}{2y}\right\} dy. \tag{4.4}$$

It will next be shown that the dynamical problem is represented implicitly by the variational equation $\delta S = 0$. Writing the integrand for short as $f(\psi_y, \psi, y)$, we have that the integral (4.4) is stationary for weak variations with fixed end-points (Fox 1950, §1.3; Bolza 1961, §17) if ψ satisfies the Eulerian characteristic equation

$$\frac{d}{dy}\left(\frac{\partial f}{\partial \psi_y}\right) - \frac{\partial f}{\partial \psi} = 0,$$

whose explicit form is

$$\psi_{yy} - H'(\psi) + \tfrac{1}{2}I'(\psi)/y = 0. \tag{4.5}$$

(Here the accents denote ψ-derivatives.) But (4.5) is identifiable with the general vorticity equation which also describes the dynamical problem completely; specifically, it is the same as equation (A 12) in the Appendix with the x-derivative put equal to zero. [Note that (4.5) can also be obtained by differentiating (4.1) and (4.2), and then using the equation of radial equilibrium $p_y = \frac{1}{2}V^2/y$ (cf. Appendix, equation (A 17)).]

We have thus established that if a curve drawn in the plane (ψ, y) is to represent a physically realizable cylindrical flow, then it must be an extremal of the integral S. It should be noted that the variations $\delta\psi$ which have to be considered in respect of the equation $\delta S = 0$ are not implied to have any physical significance; we have merely demonstrated an analytical property of any function $\psi(y)$ which describes a real flow.

Suppose a certain state of flow is fully specified, with I and H given in the first place as functions of y in the interval $(0, a)$ but hence expressible as functions of the known stream-function, say ψ_A. Then, of course, ψ_A will be a 'particular integral' of (4.5); more precisely, the known relationship of ψ_A to I and H will fix the particular form of the equation.† But (4.5) is a second-order equation whose general solution involves two arbitrary constants. Furthermore, if the total flow rate $Q = 2\pi\psi_A(a)$ is fixed, there are just two boundary conditions to be satisfied for the solution to represent another flow realizable within the same cylindrical space as the original one, namely the kinematical conditions $\psi(0) = 0$ and $\psi(a) = \psi_A(a)$ [cf. Appendix, end of §*b*]. When (4.5) is non-linear, which is the usual case, both arbitrary constants have to be chosen precisely to make the general solution satisfy these conditions (the different requirement in the exceptional case where (4.5) is linear will be explained below); and with a fair degree of generality, though allowing it may not be complete, we may assume that this can be done at least once. Thus it appears that there can exist a stream-function ψ_B which is distinct from ψ_A, which satisfies the dynamical conditions prescribed by the original flow, and which coincides with ψ_A at the end-points of $(0, a)$, so satisfying a continuity relationship with the original flow. One sees therefore that a flow of the present type can be one of a 'conjugate' pair, the second of which can be connected to the first within the same cylindrical boundary; as the sufficient attribute of the pair in respect of dynamical conditions, each stream-surface present in both flows conserves mutually the same circulation and total head in passing from one to the other. Figure 3 illustrates the possible form of a conjugate pair of solutions ψ_A and ψ_B.

It is generally not true, however, that only a single conjugate state exists for a given cylindrical flow with swirl; for instance, the two specific cases treated in

† This idea may perhaps be more easily understood by reference to the examples considered in § 5; for instance, a glance over equations (5.1) to (5.4) will show at once what is implied by the general statement above. It should be noted, however, that an *explicit* representation of the terms $H'(\psi)$ and $I'(\psi)$ in (4.5) would be impossible in many examples; that is to say, even when these terms and W are known functions of y, it might be impossible to obtain explicit forms for $\psi(y) = \int W\,dy$ and *a fortiori* for its inverse $y(\psi)$ and hence $H'(\psi)$ and $I'(\psi)$.

§5 reveal an indefinitely large number of mutually conjugate states. But vortex breakdown will be explained presently as a transition between a certain pair of 'adjacent' states, and any others indicated by the theory apparently have no physical significance. To fix the concept of these 'adjacent' states tentatively for the time being, the following observations are made. A multiplicity of conjugate states may be ordered according to their respective values of S; and, simply from the form of the integral (4.4), it is to be expected that S will get progressively larger as the corresponding solution curves joining the points [0, 0]

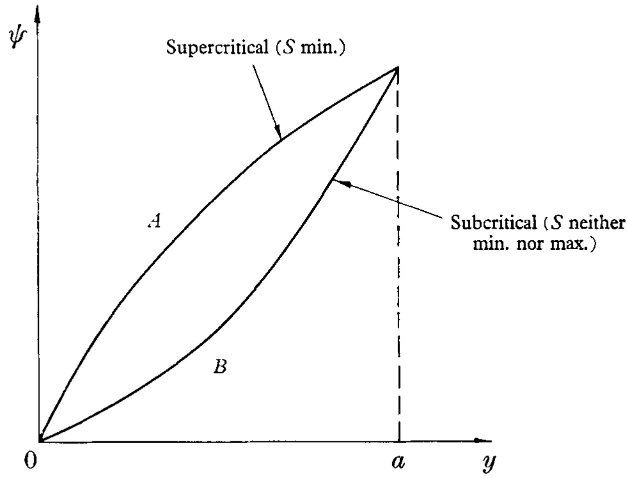

FIGURE 3. Illustration of conjugate solutions.

and $[\psi_A(a), a]$ in the (ψ, y)-plane get more sinuous. Thus there is no reason to suppose in general that the sequence of values of S would have an upper limit, but obviously it must have a lower limit. The two adjacent states in question are in fact those giving the two smallest values of S. A further interpretation is that the adjacent states are represented by solution curves which, like those in figure 3, do not intersect at their end-points.

[Before proceeding with the main argument, we need to acknowledge an exception to the present conclusions which arises when (4.5) is *linear*. For this case it is necessary that throughout $(0, a)$ both H and I take the form $k + l\psi + m\psi^2$, where k, l and m are constants, and so the physical possibilities are easily recognized. No such case is even remotely relevant as a model for the vortex breakdown phenomenon; and it is perhaps worth emphasizing that a 'combined vortex' (i.e. a core with solid-body rotation surrounded by a potential flow—see §5) is *not* a case in question, even though H and I have this particular form within each of the two distinct regions of the flow. Together with the boundary conditions, a linear form of (4.5) poses an eigenvalue problem akin to the one defined in §3, except that a small disturbance is no longer implied. If any conjugate solution is possible, then an infinitely variable set of solutions must be possible, since one constant of the general solution is left undetermined by the boundary conditions. The fact that the integral (4.4) for S then possesses a continuous field of extremals all passing through the same end-points implies that its value is the same for every one of the possible solutions.]

4.2. *Application of the variational principle*

Except in respect of the minor point covered by the preceding sentence, our variational principle is immaterial to what has been said about the existence of conjugate flows; but the principle appears to be the only means of proving several important general properties which are essential to the explanation for vortex breakdown. To this end the second variation of the integral (4.4) for S has to be considered. But before any mathematical argument relating to the second variation is introduced, it is desirable to emphasize further that the role of (4.5) as the vorticity equation is fundamental to the dynamical problem, rather than its incidental role as Euler's equation for $\delta S = 0$. For this reason the variational principle can be framed *arbitrarily* in the most convenient manner possible, namely, under the analytical restriction that a solution of (4.5) should furnish merely a weak extremum of S, and this avoids the considerable mathematical complications which relate to the generalized sufficient conditions for an extremum (Bolza 1961, §17).

For the moment leaving aside the question of sufficient conditions, we first observe that, since $\partial^2 f/\partial\psi_y^2 = 1$ is positive, Legendre's necessary condition for a *minimum* is satisfied (Fox 1950, §1.5; Bolza 1961, §11). Thus, for arbitrary weak variations about an extremal, $\psi = \psi_A(y)$ say, which is a solution of (4.5) and therefore gives $\delta S = 0$, it is possible that the second variation $\delta^2 S > 0$ so that S_A is a proper minimum, but there cannot in any case be a maximum. It appears extremely likely, furthermore, that among a set of conjugate solutions there is generally one giving a minimum of S; for we reason intuitively that the integrand of (4.4) is a positive function of ψ_y and is otherwise a finite function of position in the (ψ, y)-plane, so that between any two fixed end-points a path can presumably be found which comprises between the requirements of keeping the square of its slope small and keeping to 'low' contours of the function

$$H(\psi) - \tfrac{1}{2}I(\psi)/y.$$

To proceed on a definite basis, however, we shall assume that S is in fact a minimum for the particular extremal A representing the given flow which defines $H(\psi)$ and $I(\psi)$.

We now consider a conjugate state B which is *adjacent* to A according to the following definition. From either of the end-points in the (ψ, y)-plane, say from $[0, 0]$ for definiteness (see figure 4), an infinite set of solution curves Γ for (4.5) can be drawn with varying initial slope, thus generating a field of extremals. The curves Γ which are infinitesimally displaced from A cannot intersect it between the given end-points, for if they did they would, being themselves extremals, provide a form of variation from A for which $\delta^2 S = 0$ (see §§4.4, 4.5), so that S_A would not be a proper minimum. Thus, as their initial slope is varied from that of A, say by amount Δ, the curves first diverge away from the given arc of A as indicated by figure 4. But eventually, when $|\Delta|$ is large enough, one of the set Γ must pass through the opposite end-point of A if there is to be a conjugate solution; and it is such a curve with the least possible value of $|\Delta|$ that we define as the adjacent conjugate B. We have therefore that B intersects A only at the end-points, and that extremals Γ which are contained between A and B all intersect B but none intersects A.

4.3. *Preliminary demonstration of properties of conjugate flow B*

We go on to establish as general properties that $S_B > S_A$, and that the integral giving S_B is neither a minimum nor maximum. In the first place a simple partially geometrical argument will be presented; but later the conclusions will be confirmed by a more powerful analytical demonstration.

Suppose that the space between A and B is covered by an infinite set of curves $\psi = \eta(y)$ which may depend on one or more parameters and which connect the end-points without anywhere else overlapping. Obviously none of these curves can be a single extremal over its entire length, except the two bounding the set by coinciding with A and B. We now need to consider whether it is possible to construct the curves so that the quantity S_η, found by performing the integration of (4.4) along $\psi = \eta$, varies monotonically through the set. For successive curves $\psi = \eta$ and $\psi = \eta + \delta\eta$, the variation of S_η is, by an elementary result of the calculus of variations,

$$\delta S_\eta = 2\pi\rho \int_0^a L\delta\eta\, dy, \tag{4.6}$$

where

$$L = H'(\eta) - \tfrac{1}{2}I'(\eta)/y - \eta_{yy}. \tag{4.7}$$

Since $\delta\eta$ always has the same sign, our proposition will be established if we show that L can have the same sign at all points along every curve of the set. Note that $L = 0$ is Euler's equation (4.5), so that the integrand of (4.6) will vanish wherever $\psi = \eta$ coincides with an extremal over a finite arc. (For instance, δS_η becomes a quantity of second-order smallness as $\eta \to \psi_A$, which makes $L \to 0$.) It also follows that at any point $L < 0$ when $\eta_{yy} > (\psi_e)_{yy}$ and vice versa, if $\psi = \psi_e$ be an extremal pasing through this point.

When the set of curves is begun at A, and B lies below as in figure 4, then we have $\delta\eta < 0$. Since S_A is a minimum, δS_η must become positive as η begins to depart from ψ_A; and so a choice of curves giving $L \leqslant 0$ is a possibility, whereas $L \geqslant 0$ is impossible. (The equality sign in $L \leqslant 0$ means only that L may vanish over *part* of a particular η-curve; it has already been made clear that A and B are the only curves with $L = 0$ everywhere.) Thus it remains to show that a choice can be made giving $L \leqslant 0$ and hence $\delta S_\eta > 0$ everywhere. One intuitively sees the feasibility of this as soon as figure 4 is reconsidered, remembering that the field of extremals Γ radiating from the origin between A and B stays below A and crosses only B. Consequently, any η-curve starting from the origin in alignment with one of Γ has to be curved positively away from this extremal, so making $L < 0$, in order to be taken to the opposite end-point. Moreover, it appears that the η-curves can always be steered through the part of Γ initially above them in such a way as to keep $L \leqslant 0$. Hence we conclude that a set of η-curves can in fact be chosen so as to make S *increase steadily* as B is approached from A; and a corresponding argument with obvious modifications establishes the same conclusion for the case where B lies above A. (But note that this conclusion would certainly not hold if there were any extremal between A and B which also passed through the given end-points.) By the existence of any such choice of η-curves, it follows immediately that $S_B > S_A$ and that B cannot give another minimum; and, as explained earlier, B cannot give a maximum.

There is one particular choice for the η-curves which deserves noting. Suppose *each curve follows an extremal* Γ *up to near its intersection with* B, *then turns* rapidly and continues to the opposite end-point along B. We then have $L = 0$ along each arc of the curve coincident with the respective extremals Γ and B, and the integral (4.6) is comprised entirely from the contribution at the corner between the two arcs. When corners of small yet finite size are considered, it appears very clearly that $\delta S_\eta > 0$; but the particular expression (4.6) for δS_η is not directly determinate in the limit as the corner size is reduced to zero. However, *the value of* δS_η *in this limit is immediately* forthcoming from Weierstrass's theorem, which will be used in §4.5 to obtain an expression for $S_B - S_A$. Note, by the way, that in the limiting case of the present η-curves the sharp corners occur over only part of B; for their abscissae have a lower limt y_1, at which B first enters the field Γ (this is the abscissa of the 'conjugate point' or 'kinetic focus' of B, which is to be discussed in §4.4).

On further consideration of figure 4 it can be inferred that a solution curve for (4.5) started from the origin with a slope somewhat less than the slope of B will intersect B for $y < y_1$. Hence we see the possibility of a third conjugate solution such as represented by curve C in the figure; the fact that C also intersects A just once between the end-points can be shown by reasoning on the same lines as hitherto, but applied to a field of extremals radiating from the end-point $[\psi_A(a), a]$. To compare S_C with S_B, the two separate regions circumscribed by C and B may be considered in turn, and for each the variation of S_η examined for η-curves drawn between one end-point and the intermediate point of intersection between C and B. The conclusion is that each of the two contributions to S_C exceeds the respective contribution to S_B, so that $S_C > S_B$, and S_C also is not a minimum. By a continuation of this form of argument, it can be concluded that further conjugate solutions must give still larger values of S, none of which can be a minimum. We cover this matter only briefly here because evidently the existence of conjugate solutions other than A and B is not relevant to the physical problem. It may be noted, however, that the property of successively larger values of S for multiple conjugates is shown very clearly by example 2 in §5.

4.4. *Interpretation of supercritical and subcritical states*

On recalling a fundamental result from Jacobi's theory of the second variation (Fox 1950, Ch. 2; Bolza 1961, Ch. 2), we discover an important physical implication of the properties so far established. We have to consider particular variations $\epsilon\phi(y)$ about an extremal, say $\psi = \psi_A$, whose arc over the interval $(0, a)$ is to be tested for a minimum of S; here ϵ is a constant of first-order smallness, and ϕ is a solution of Jacobi's 'accessory equation'. Because

$$\partial^2 f/\partial \psi_y^2 = 1 \quad \text{and} \quad \partial^2 f/\partial\psi\,\partial\psi_y = 0$$

in the present instance, the accessory equation is given simply by putting $\psi = \psi_A + \epsilon\phi$ in Euler's equation (4.5) and linearizing in ϵ; thus it is

$$\phi_{yy} - \{H''(\psi_A) - \tfrac{1}{2}I''(\psi_A)/y\}\,\phi = 0. \tag{4.8}$$

Let ϕ be constructed from the general solution of (4.8) so that $\phi = 0$ at one end-point of the integral S (i.e. $\phi(0) = 0$ or $\phi(a) = 0$). Then, respective to the end-point taken, the 'conjugate point' or 'kinetic focus'† is by definition the nearest point on the extremal whose abscissa y_1 makes $\phi(y_1) = 0$. If it lies beyond the opposite end-point of the integral, this and Legendre's condition are sufficient for a minimum. If it lies nearer than the opposite end-point, however, then S is neither a minimum nor a maximum irrespective of Legendre's condition. Hence we see that each of these cases is represented in a conjugate pair of flow states A and B as previously defined; that is, the extremal arc A has no 'kinetic focus' between the end-points, but B has one. Furthermore, extremal arcs C, etc., giving further conjugate states are all characterized in the second way.

Now, the vital point here is that the accessory equation (4.8) is the same as (3.3), which was introduced specifically as the equation satisfied by the 'test function' ϕ_c, but which is basically the dynamical equation for infinitesimal standing waves of extreme length. The equivalence of the two equations is demonstrated in the Appendix, §*b*, where the coefficient of ϕ in (4.8) is worked out explicitly as a function of y. Hence we recognize the criterion for a minimum of S to be precisely equivalent to the test for supercritical flow conditions that was explained in §3. Again, subcritical conditions are seen to imply the existence of a 'kinetic focus' cancelling the minimizing property of the extremal arc which represents the flow.

4.5. *Further analytical significance of conjugate-flow properties*

The arguments used in §4.3 are to an extent intuitive, and so it is desirable to supplement them with the following analytical interpretation of the main facts which have been revealed about the adjacent flow states A and B. First, however, we should note that the rather difficult question of the existence in general of a field of extremals need not concern us here; for it is satisfactory to regard the set of solution curves Γ as being implicit in the general physical model. That is to say, the existence of Γ is a necessary attribute to any system specified realistically by a particular choice of forms for $I(\psi)$ and $H(\psi)$. The *physical* necessity of the field can be recognized in the need for a range of solutions representing cases where the flow is connected into a second cylindrical space bounded at $y \neq a$. Again, a continuous set of solution curves radiating inward from the opposite end-point is necessary to represent the range of physically possible cases where the flow is connected into an annular space.

The curves Γ with the origin as their common point may be defined by the expression

$$\psi = \overline{\psi}(y; k), \tag{4.9}$$

where k is a parameter—which could be, for instance, the initial slope of the solution $\overline{\psi}$. Obviously, we have $\overline{\psi}_k (\equiv \partial\overline{\psi}/\partial k) = 0$ at $y = 0$ over the complete range of k. For any particular extremal of this set, say one specified by $k = k_0$,

† The former term is far more common, and the latter is strictly appropriate only to problems of particle trajectories in which there is a special mechanical interpretation of conjugate points; but the latter term will be used hereafter to obviate any confusion with our extensive other use of the adjective 'conjugate'.

it can be shown that its 'kinetic focus' is at $y = y_1$, where y_1 is the next root above zero of the equation

$$\psi_k(y;\, k_0) = 0; \tag{4.10}$$

and this also defines the point of contact of the curve with the *envelope* of the set Γ (Fox, §2.9; Bolza, §15).

Accordingly we arrive at the interpretation indicated by figure 4, in which the envelope is drawn as a dashed line (see also figure 6). Recognizing A and B to belong to a set of curves with a common envelope as shown in this figure, we

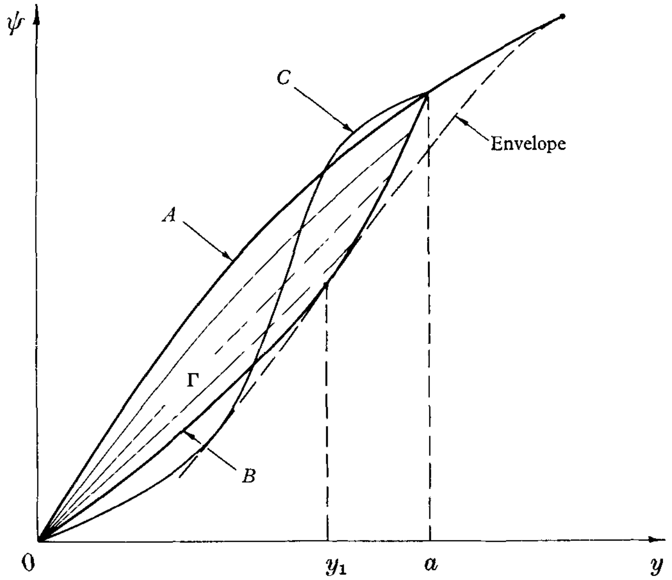

FIGURE 4. Guide to the analytical interpretation of conjugate-flow properties.

see at once that B must touch the envelope for $y < a$ if A does so for $y > a$, whereupon our identification of the envelope with the locus of the kinetic foci establishes the flow states A and B to be respectively supercritical and subcritical. It should be remembered, however, that the physical problem as basically posed gives the extremal arc A only over the interval $(0, a)$, so that to produce A beyond $y = a$ an analytic continuation of $I(\psi)$ and $H(\psi)$ is necessary beyond the range of ψ in which they are first defined. Thus, no essential physical process is represented by the continuation of A, or in general by any construction in that part of the (ψ, y)-plane outside the strip $0 \leqslant y$, $0 \leqslant \psi \leqslant \psi_A(a)$; but of course those details of figure 4 lying outside this 'physical domain' are nevertheless wholly determined by the specifications of the physical problem.

To demonstrate that $S_B > S_A$ in general, use may be made of Weierstrass's theorem (Fox, §9.2; Bolza, §20). Thus we have, for the total variation of S between A and B,

$$S_B - S_A = 2\pi\rho \int_0^a \mathbf{E}(p_e, p_B;\, \psi_B, y)\, dy, \tag{4.11}$$

where the Weierstrass $\mathbf{E}$-function has now to be evaluated in respect of B, whose slope is denoted by p_B, passing through the field of extremals Γ, whose slopes at points $[\psi_B, y]$ are denoted by p_e. Recalling the form of the integrand in (4.4) we get

$$\begin{aligned}\mathbf{E} &= f(p_B, \psi_B, y) - f(p_e, \psi_B, y) + (p_e - p_B) f_p(p_e, \psi_B, y) \\ &= \tfrac{1}{2}(p_B^2 - p_e^2) + (p_e - p_B) p_e = \tfrac{1}{2}(p_B - p_e)^2. \end{aligned} \tag{4.12}$$

[Note that for η-curves of the type with sharp corners defined in the penultimate paragraph of §4.3, the variation δS_η is given by $\mathbf{E}\delta y$, if δy is the difference in the abscissae of the corners of successive curves.] The result obtained from (4.11) and (4.12) is

$$S_B - S_A = \pi\rho \int_{y_1}^{a} (p_B - p_e)^2 \, dy, \tag{4.13}$$

where the lower limit of integration has been replaced by y_1 because obviously $\mathbf{E} = 0$ in the interval $(0, y_1)$. Thus we find that $S_B - S_A$ is positive, as otherwise demonstrated in §4.3.

4.6. *The explanation for vortex breakdown*

It has now been proved generally that, for a primary flow which is supercritical in the sense defined in §3, the conjugate state is subcritical and has a larger flow-force S. Together with the facts, noted at the end of §3, that the periodic waves which can be superposed on a subcritical flow cause a reduction in S and that they form only downstream from their originating agency, these results demonstrate a set of properties comprising a complete analogy with the set which was noted in §2 for open-channel flows, and which supported an entirely adequate explanation for the hydraulic jump. Our explanation for the vortex phenomenon is virtually complete, therefore, and for the details of the physical argument we need only refer back to §2.

To sum up the essentials, vortex breakdown is explained as a 'finite-amplitude' transition from a supercritical state of flow upon which infinitesimal standing waves cannot occur. The conjugate subcritical state has an excess of flow-force which, if the flow were first held steady after the transition and then released, would be bound to generate an unsteady motion. For a mild transition, the excess of flow-force is small enough to be cancelled by the formation of waves in the subcritical régime created downstream, without necessitating significant energy loss. But for a strong transition the effect of wave resistance is inadequate to establish the flow-force balance demanded by a steady state; an energy loss is then required, which means that in practice a region of vigorous turbulence is generated. (Note the clear indication given by (4.4) that S is reduced by a decrease in the total head H.)

The following practical aspect comes to light when specific cases are considered. It appears that in examples typical of swirling flows in practice, the curve B representing the subcritical conjugate state in the (ψ, y)-plane generally lies below the curve A representing the supercritical primary flow, i.e. as shown in figures 3 and 4. For instance, although not forthcoming in the first example treated in §5 below, this property is found in the much more realistic second example. Thus, a tendency towards stagnation on the axis is apparently to be associated with the breakdown phenomenon, as indeed is indicated by the experimental facts (e.g. see Harvey 1962).

Another practical aspect worth emphasizing, recalling §3.3, is that supercritical states are, in a general classification amongst possible flows, characterized by comparatively small amounts of swirl. That this must be so becomes fairly obvious on recognition that swirl provides the 'stiffness' element analogous to gravity in the hydraulic problem—more precisely, it does so where the square of the circulation increases outwards, as usual near the axis of a real vortex. This general role of swirl is indicated clearly by equation (3.1); we see that a larger (positive) I_y always gives a more rapid rate of oscillation of the solution and hence a greater possibility of an eigenvalue $\gamma^2 = -\alpha^2 < 0$ representative of standing waves. Zero swirl, the most innocuous case in present respects, corresponds to an indefinitely far supercritical state, just as zero gravity corresponds to an infinite Froude number.†

4.7. *Special cases of conjugate flow B*

It needs to be emphasized that the subcritical cylindrical flow B cannot occur *intact* in the rear of a vortex breakdown, even though its existence as a conjugate to the primary flow A is an essential element of the breakdown mechanism. Nevertheless, for a mild breakdown manifesting a wave structure, the actual flow is very likely to preserve the main features of B; that is, the theoretical axial-velocity and swirl-velocity distributions for B may still be recognizable more or less distinctly despite the deformation into waves. We shall now consider two instances of a specially distinctive feature of an actual flow which may be represented by the solution for B.

Particularly in cases of a mild breakdown, a region of reversed flow is sometimes observed near the axis (see Harvey 1962); and this might correspond approximately to a theoretical case such as is illustrated in figure 5(*a*). Here the portion of the curve B lying below the y-axis, for $0 < y < y_c$, represents an eddy of fluid contained within the stream-surface $\psi(y_c) = 0$ which connects to the axis in the primary flow upstream from the transition. It must be noted, however, that this inner flow is simply an analytic continuation of the main flow (i.e. with $0 \leqslant \psi \leqslant \psi_A(a)$) which originates upsteam, and so it has no essential correspondence with the actual physical circumstances of a central eddy—in which the viscosity of the fluid is inevitably an important factor.

Another anomaly of this type of solution B is that apparently the flow may be unstable. In keeping with this interpretation, the means explained in the Appendix, §*c*, for finding the group velocity of standing waves becomes indeterminate when $W = \psi_y$ changes sign in $(0, a)$; although a standing-wave solution of the small-disturbance equation can be found, there appears to be no neighbouring travelling-wave solution with a real value of phase velocity. Instability of the subcritical régime is evidently a common feature of vortex breakdown in practice; but, for the reason noted at the end of the last paragraph, this aspect of the theory is unlikely to correspond precisely with an actual instability.

† At first sight this matter tends to be somewhat obscured by one's natural inclination to regard very high Froude numbers as 'dangerous'. The slight conceptual difficulty which may confront the practically minded arises from the fact that in hydraulic engineering a high Froude number means a large flow velocity, never a small g!

A related case which has a special application free from the previous difficulties is illustrated in figure 5(*b*). Here the flow B surrounds a core of stagnant fluid, or a cavity, which extends from the axis to $y = y_c$. Now, although this is an unrealistic model for a vortex breakdown, it well describes the situation where a filamentary vapour cavity is formed along a vortex in water, the best-known instance being the cavitation of tip vortices shed from ships' propellors.

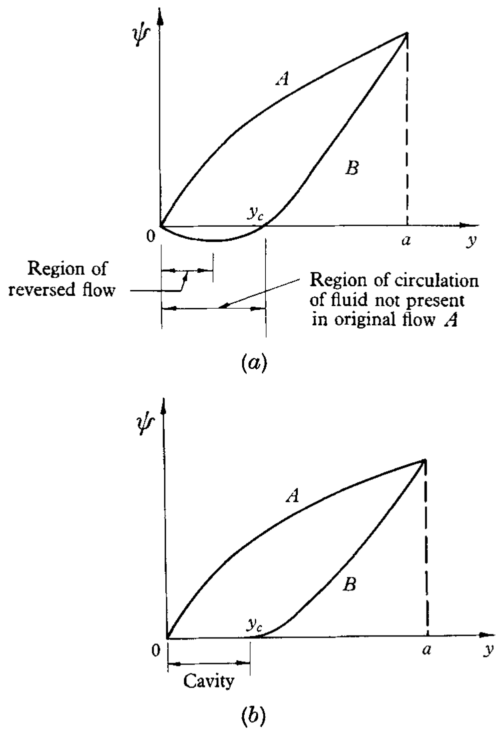

FIGURE 5. Special cases of conjugate flow B.

The latter phenomenon is not, of course, a breakdown in the sense we have defined; for irrespective of the previously considered factors the specific condition of formation is that the pressure on the axis—which gradually decreases in the flow direction as the vortex trailing from a blade rolls up—ultimately falls to vapour pressure. However, since the beginning of the cavity is evidently describable as a transition between two conjugate (energy-conserving) flows, we can *ipso facto* apply some of the ideas of this paper. Indeed, they appear to be necessary to show in general why such a cavity has standing waves upon it, as is observed experimentally. We assume that the flow A preceding the cavity is supercritical, which is reasonable since cavitation would be less likely to occur after an ordinary breakdown to a subcritical state (which would reduce the velocity on the axis and so raise the pressure—cf. §5, example 2). This assumption is anyway bound to hold in some circumstances because, by adjustment of the environmental pressure level, the cavitation condition can be produced independently of the supercritical-flow condition.

The solution B is no longer a conjugate of A according to the fundamental definition since the respective extremal curves now have only one end-point in common, and so the general conclusions hitherto established do not apply directly. The essential situation is found, however, to be just the same as before.

First, it appears that the flow B is always subcritical in the general sense, because standing waves can always occur in such a flow surrounding a cavity at constant pressure. But, for obvious reasons, this fact no longer implies the existence of a 'kinetic focus' at an interior point of the extremal arc B. For the following argument we shall assume for simplicity that it is actually a minimizing arc like A; but the case where it is not requires only a minor complication of the argument. According to this assumption, the curves A and B can be considered as members of a field of extremals Γ' radiating from their common end-point at $y = a$, those of which between A and B cross the y-axis in the interval $(0, y_c)$. The slope of an extremal at its point of crossing the y-axis will be denoted by s.

The flow-force excess for the subcritical régime is now given by

$$\Delta S = S_B + 2\pi p_0 y_c - S_A, \tag{4.14}$$

where p_0 is the pressure within the cavity. In the physical situation of a cavitating vortex, the cavity must have a stagnation point at its forward end, which means that $p_0 = \rho H_0$, where H_0 is the value of the total head on the axis. The integrand of (4.4) reduces to H_0 if the path of integration in the (ψ, y)-plane is taken along the y-axis; and therefore ΔS is the total variation of the integral between the curve A and the curve comprising B together with the portion of the y-axis from the origin to y_c. Hence Weierstrass's theorem used just as in §4.5, but now with respect to the field Γ', shows directly that

$$\Delta S = \pi\rho \int_0^{y_c} s^2\, dy. \tag{4.15}$$

Thus, as was expected, ΔS is necessarily positive.

Note that the existence of a stagnation point on the surface of a cavity at constant pressure implies that the fluid is at rest everywhere on its surface. The boundary condition $W(y_c) = 0$, together with $\psi(y_c) = 0$ and the kinematical condition at $y = a$, will be enough to determine a conjugate solution ψ_B from the general solution of (4.5).

4.8. *Concluding theoretical points*

Two matters remain deserving attention. There is first the question of whether or not our general conclusions depend significantly on the feature of our theoretical model that the flow is finitely bounded in the radial direction. In the case $a \to \infty$, perhaps the main innovation arising in the theory concerns the general eigenvalue problem explained in §3. Whereas negative eigenvalues $\gamma^2 = -\alpha^2$ still occur as discrete numbers, if at all, the system now admits a continuous spectrum of positive ones. As before, however, only the negative ones have a fundamental role in our physical interpretation of flow properties, and so it would appear that the general conclusions respecting supercritical and subcritical states still stand. Again, extremal curves of infinite length pose considerable analytical difficulties, but intuition suggests that the leading deductions from our variational principle should still remain valid. As evidence of this, example 2 in §5 gives for an unbounded flow clearly determinate results which are wholly consistent with the general theory.

The second point concerns the assumption of axial symmetry. As Harvey (1962) has shown, the vortex breakdown phenomenon can be made almost exactly axisymmetric; and though under less rigorous experimental conditions the phenomenon may be subject to a considerable degree of asymmetry, and also to fluctuations in time, there seems little doubt that its essential mechanism is explained by the present model of a steady axisymmetric flow (just as well, one might say, as hydraulic jumps are explained essentially by a two-dimensional model, notwithstanding that they often occur in practice with considerable spanwise non-uniformity). From a theoretical viewpoint, however, the question naturally arises whether there exists analogous mechanisms with different overall geometry. Evidently the answer is negative, as may be reasoned with reference to a general form of infinitesimal standing waves whose dependence on the cylindrical polar angle θ is $\cos n\theta$. The wave properties for the present case $n = 0$ are found to be essentially different from those for $n > 0$, so that our arguments regarding supercritical and subcritical states break down in the other cases. For a stable cylindrical flow, it appears that standing waves with $n = 1$ first become possible at very *small* wavelengths as the swirl is increased from zero; thus the system is less 'stiff' for longer waves, the opposite of the case $n = 0$. For $n > 1$, standing waves are always possible whatever the (non-zero) magnitude of the swirl, and so no supercritical state can be classified.

5. Examples

Two examples will be given to illustrate the results which have been proved generally. The first provides a simple non-linear form of (4.5) which, though it cannot be solved explicitly, allows us to demonstrate clearly the existence of conjugate solutions; but unfortunately the distribution of swirl velocity assumed for this example is not typical of real flows. The second example is the only one which has been found admitting solution in closed form, but it presents an exceptional aspect of the general analytical problem. This 'combined vortex' flow consists of two parts for each of which equation (4.5) takes a linear form, and the 'non-linearity' essential† to the overall problem enters through the intermediary boundary conditions, which apply on a stream-surface specified by the dependent variable ψ rather than the independent variable y. As ψ and $W = d\psi/dy$ are made continuous across this interface, as also are the coefficients of the respective linear equations (though the derivatives of the coefficients are not), the system provides the same overall properties as a wholly continuous second-order non-linear system of the general class considered in §4.

† Essential to avoid degeneracy of the sort explained in §4 with reference to linear systems.

Example 1. *Simple illustration of non-linear problem*

For the primary flow A in $0 \leqslant y \leqslant a$, we take

$$\left.\begin{aligned} W &= 1 \quad \text{so that} \quad \psi = y, \\ \text{and} \qquad V &= (\tfrac{2}{3})^{\frac{1}{2}}\,\kappa y, \end{aligned}\right\} \tag{5.1}$$

where κ is a constant. Hence we have

$$I = yV^2 = \tfrac{2}{3}\kappa^2 y^3 \equiv \tfrac{2}{3}\kappa^2\psi^3,$$

so that

$$I'(\psi) = 2\kappa^2\psi^2. \tag{5.2}$$

And, by equation (A 18) in the appendix, we have

$$H'(\psi) = \kappa^2 y \equiv \kappa^2\psi. \tag{5.3}$$

The general equation (4.5) for the stream-function therefore becomes

$$\psi_{yy} = \kappa^2\left(\psi - \frac{\psi^2}{y}\right). \tag{5.4}$$

A particular solution of (5.4) obviously is $\psi = \psi_A = y$, representing the primary flow.

We shall now demonstrate the existence of second solutions ψ_B which satisfy boundary conditions $\psi_B(0) = 0$ and $\psi_B(a) = \psi_A(a) = a$. In the $(\psi,\, y)$-plane, consider a solution curve B started from the origin with a positive slope less than unity, so that initially B lies below curve A, i.e. below $\psi = y$ (see figure 6). Equation (5.4) shows that $\psi_{yy} > 0$ for $0 < \psi < y$, and so B has positive curvature wherever it lies below A. Hence B must inevitably intersect A if sufficiently far produced. Again, if B is started from the origin with a slope greater than unity, its curvature is everywhere negative while it remains above A, so that it is bound ultimately to intersect A away from the origin. [The two B-curves in figure 6 are drawn accurately for $\kappa = 1$, their initial slopes being $\frac{3}{2}$ and $\frac{1}{2}$. The curves intersect $\psi = y$ at $y = a_1 = 2{\cdot}76$ and $y = a_2 = 3{\cdot}81$, respectively.]

Note incidentally that $\psi_{yy} < 0$ for $\psi < 0$, which proves that a solution curve started from the origin with a negative slope cannot again intersect $\psi = y$. Thus a conjugate solution with a region of reversed flow is impossible in this example.

Since the point of intersection of B with A will obviously vary with the choice of positive initial slope for B, we may conclude that conjugate solutions exist for a range of problems, i.e. with different a. Next we have to consider the nature of possible conjugates for a given a.

First, the condition for critical flow is derived. Putting $\psi = y + \phi$ in (5.4), we get

$$\phi_{yy} + \kappa^2\phi = -\kappa^2\phi^2/y, \tag{5.5}$$

whose linearized form

$$\phi_{yy} + \kappa^2\phi = 0 \tag{5.6}$$

is evidently the equation for long standing waves also interpretable as Jacobi's accessory equation for the extremal $\psi_A = y$. The solution of (5.6) vanishing at $y = 0$ is

$$\phi = C\sin\kappa y, \tag{5.7}$$

whose next zero is at $y = \pi/\kappa$. Hence, recalling the results of §§3 and 4, we conclude that the primary flow is supercritical (curve A giving a minimum of S) if $a < \pi/\kappa$, and subcritical if $a > \pi/\kappa$.

Now, comparison of (5.5) and (5.6) shows that for a given ϕ (i.e. for a given displacement from curve A), the value of ϕ_{yy} for the solution of the linearized equation is greater by amount $\kappa^2\phi^2/y$ than the value for the 'exact' solution. From this it follows that a solution curve B started with positive slope less than unity (e.g.

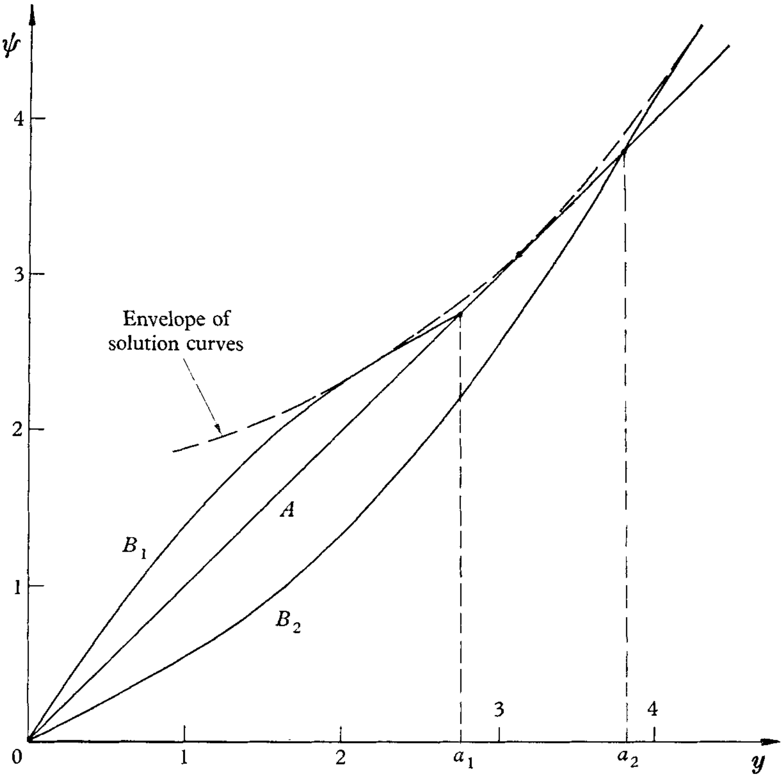

FIGURE 6. Solutions of $\psi_{yy} = \psi - \psi^2/y$.

curve B_2 in figure 6) will not intersect A until beyond $y = \pi/\kappa$ where the solution of (5.6) has a zero. Again, a solution curve like B_1 in figure 6 will necessarily intersect A for $y < \pi/\kappa$. Thus we have that a supercritical primary flow possesses a conjugate of type B_1 with increased velocity near the axis (this is an unrealistic property which may be regarded as a consequence of the unrealistic assumption that $V \propto r^2$). We also have that a solution curve like B_2 can only represent a conjugate to a subcritical case of flow A.

It remains to verify that a conjugate flow of type B_1 is always subcritical, and one of type B_2 always supercritical. To do this the accessory equation for an extremal $\psi = \psi_B(y)$ has to be considered. Putting $\psi = \psi_B + u$ in (5.4), we obtain

$$u_{yy} + \kappa^2\left(\frac{2\psi_B}{y} - 1\right)u = -\frac{\kappa^2 u^2}{y}, \tag{5.8}$$

whose linearized form is the accessory equation. For a case like B_1, comparison of the coefficients in (5.6) and (5.8) shows the solution of the accessory equation for B to oscillate more rapidly than the solution (5.7) for A; but, of course, this observation is inadequate to prove that the present solution oscillates inside the interval $(0, a)$. However, we know there is an exact solution $u = y - \psi_B$ of the non-linear equation (5.8); and this solution, representing the displacement of curve A from curve B, oscillates exactly once over $(0, a)$. By an argument similar to the one used in the last paragraph, we deduce that in the present case of B_1 this finite solution oscillates less rapidly than the solution of the linearized equation obtained from (5.8); and it follows immediately that B_1 has a 'kinetic focus' inside $(0, a)$, so that a subcritical flow is represented. A corresponding argument just as readily proves that B_2 must represent a supercritical flow.

As an implication of these results, it may be considered that the envelope of a field of extremals radiating from the origin touches $\psi = y$ at $y = \pi/\kappa$ and has positive curvature there. This envelope is sketched as a dashed line in figure 6.

It is not true that for a supercritical primary flow the subcritical conjugate B_1 is unique, and we have to recognize the possibility of an indefinitely large number of further subcritical flow states represented by solution curves which oscillate more than once about $\psi = y$ in the interval $(0, a)$. However, these will give values of S larger than the values for the state B_1, which can be considered as naturally adjacent to the given supercritical state A. As explained earlier, the physical mechanism of vortex breakdown is to be interpreted essentially as a transition between such adjacent states, the existence of further subcritical conjugates being an irrelevant side-issue of the analysis.

For long travelling waves with phase velocity c in the x-direction (see Appendix, §d), the equation for the stream-function perturbation respective to the primary flow is

$$\phi_{yy} + \frac{\kappa^2}{(1-c)^2}\phi = 0, \tag{5.9}$$

and the boundary conditions are $\phi(0) = \phi(a) = 0$. Hence, for the mode with only one oscillation in $(0, a)$, which gives the largest wave velocity $|c-1|$ relative to the flow, we have

$$\frac{\kappa}{1-c} = \pm\frac{\pi}{a} = \pm\kappa_c, \tag{5.10}$$

where $\kappa_c = \pi/a$ is the value of κ which makes the flow critical for a given a. The two values of c given by (5.10), i.e.

$$c_+ = (\kappa/\kappa_c)+1, \quad c_- = -(\kappa/\kappa_c)+1,$$

refer to propagation with and against the flow, respectively. The characteristic number defined by (3.4) is therefore simply

$$N = \kappa_c/\kappa. \tag{5.11}$$

This result exemplifies nicely the general property that supercritical flow states ($N > 1$) are characterized, in comparison with subcritical ones, by smaller swirl velocities.

Example 2. Combined forced and free vortex

The primary flow has uniform axial velocity and consists of two parts: (I) a core with solid-body rotation, and (II) a surrounding annular region of *irrotational* vortex motion. For this example it is convenient to use the radius r as independent variable, rather than $y = \frac{1}{2}r^2$ as hitherto, and to take the core radius as length unit. Accordingly we assume

$$W = 1, \quad \psi = \tfrac{1}{2}r^2 \quad \text{in} \quad 0 \leqslant r \leqslant R, \tag{5.12}$$

and, defining the two regions of flow,

$$\left.\begin{array}{lll} \text{(I)} & V = \omega r & \text{in} \quad 0 \leqslant r \leqslant 1, \\ \text{(II)} & V = \omega/r & \text{in} \quad 1 \leqslant r \leqslant R. \end{array}\right\} \tag{5.13}$$

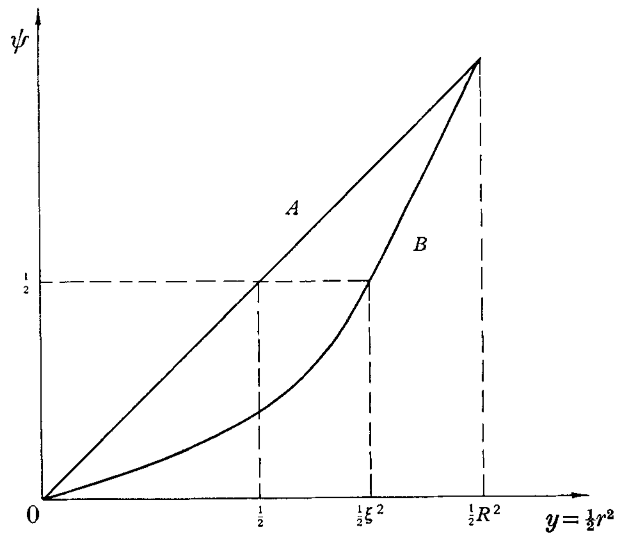

FIGURE 7. Form of the conjugate solutions for a combined vortex with finite radial boundary.

It is supposed that in a conjugate flow B the boundary between the two regions is displaced to $r = \xi$ (figure 7). At $r = \xi$ we must have $\psi = \psi_A = \frac{1}{2}$, and, since the circulation, total head and pressure are to be continuous, the axial velocity $r^{-1}(d\psi/dr)$ must be continuous.

Region II is conveniently considered first. From the assumed conditions it is readily verified that I and H are constant over this region (in fact equal to $\frac{1}{2}\omega^2$ and $H_0 + \omega^2$, respectively, where H_0 is the total head on the axis), as must be the case since the flow is irrotational. Thus equation (4.5) reduces to $\psi_{yy} = 0$, which implies $\psi_{rr} - r^{-1}\psi_r = 0$, and the solution satisfying

$$\psi(\xi) = \tfrac{1}{2} \quad \text{and} \quad \psi(R) = \psi_A(R) = \tfrac{1}{2}R^2$$

is

$$\psi(r) = \frac{(1-\xi^2)\,R^2 + (R^2-1)\,r^2}{2(R^2-\xi^2)}, \tag{5.14}$$

which shows the (constant) axial velocity in II to be

$$W_{\mathrm{II}} = \frac{R^2-1}{R^2-\xi}. \tag{5.15}$$

For region I we obtain, proceeding just as in the previous example,

$$H'(\psi) = 2\omega^2, \quad I'(\psi) = 4\omega^2\psi. \tag{5.16}$$

Hence equation (4.5) becomes

$$\psi_{rr} - r^{-1}\psi_r + 4\omega^2\psi = 2\omega^2 r^2. \tag{5.17}$$

The most general solution of (5.17) satisfying $\psi(0) = 0$ may be expressed in the form

$$\psi = \tfrac{1}{2}r^2 - \tfrac{1}{2}(Cr/\omega)\, J_1(2\omega r), \tag{5.18}$$

where C is an arbitrary constant, and the corresponding form for the axial velocity is

$$W = 1 - CJ_0(2\omega r). \tag{5.19}$$

[Note that (5.19) derives from (5.18) in consequence of the relation

$$J_1'(x) + x^{-1}J_1(x) = J_0(x)$$

between Bessel functions.]

The boundary condition $\psi(\xi) = \frac{1}{2}$ gives

$$\xi^2 - (C\xi/\omega)\, J_1(2\omega\xi) = 1, \tag{5.20}$$

and the condition $W(\xi) = W_{\mathrm{II}}$ gives

$$1 - CJ_0(2\omega\xi) = \frac{R^2-1}{R^2-\xi^2}. \tag{5.21}$$

Elimination of C between (5.20) and (5.21) leads to

$$\frac{J_0(2\omega\xi)}{J_1(2\omega\xi)} = -\frac{\xi}{\omega(R^2-\xi^2)}, \tag{5.22}$$

which is an implicit equation for ξ. When ξ is found from (5.22), C is given by either of (5.20) or (5.21), and the conjugate flow is then determined completely.

An important feature of the present example is that it still provides a clearly determinate solution when the outer boundary is expanded to infinity. The nature of the conjugate ψ_B in this case is illustrated in figure 8, which shows in particular that $W_{\mathrm{II}} = 1$, as is indicated by (5.15) and as is obviously to be expected on physical grounds. Letting $R \to \infty$ in (5.22), we get the condition $J_0(2\omega\xi) = 0$. Hence, taking the first zero of the Bessel function (which clearly is the one that gives the 'adjacent' conjugate state whose solution curve does not oscillate about the curve A), we have

$$\omega\xi = 1{\cdot}20^+ \quad \text{for} \quad R = \infty. \tag{5.23}$$

For finite R, the left-hand side of (5.22) is negative, and so $\omega\xi$ lies within the interval $1{\cdot}20^+$ to $1{\cdot}92^-$ over which $J_0(2\omega\xi)$ is negative and $J_1(2\omega\xi)$ is positive, approaching zero at the upper limit.

The condition for critical flow is next derived. Although this can be done as before by consideration of the equation for long standing waves (as was in fact done by Squire (1960) for this example), the outcome is already evident from our present results. Since a state approaching critical is characterized by the existence of a conjugate solution infinitesimally displaced from ψ_A, so that $\xi - 1$ is infinitesimal, it follows that the required limiting condition is obtained simply by putting $\xi = 1$ in (5.22). Thus, the critical value of ω is given by

$$\frac{\omega_c J_0(2\omega_c)}{J_1(2\omega_c)} = -\frac{1}{R^2-1}. \tag{5.24}$$

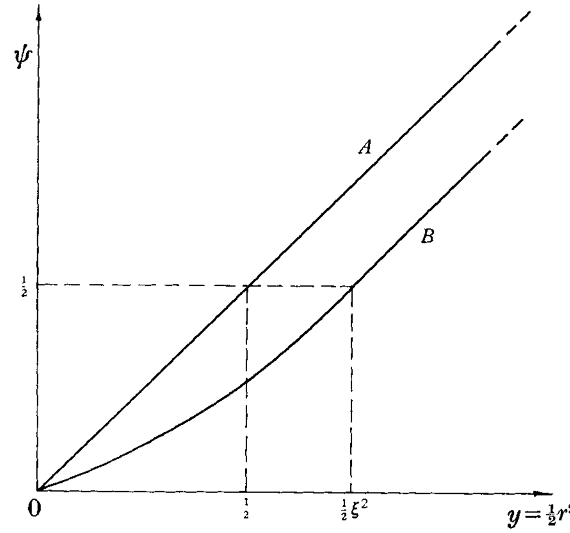

FIGURE 8. Form of the conjugate solutions for a combined vortex which is unbounded in the radial direction.

In particular we have

$$\omega_c = 1.20^+ \quad \text{for} \quad R = \infty, \tag{5.25}$$

so that (5.23) can be written

$$\xi = \omega_c/\omega \quad \text{for} \quad R = \infty. \tag{5.26}$$

For long standing waves the stream-function perturbation in I is the same as the second term on the right-hand side of (5.18), but with C made infinitesimal; and clearly this oscillates more rapidly with increasing ω. Hence we conclude that the primary flow is supercritical when $\omega < \omega_c$, and subcritical when $\omega > \omega_c$.

For a supercritical primary flow, (5.22) and (5.24) show therefore that the rotational core is expanded ($\xi > 1$) in the subcritical conjugate flow; and it follows that the axial velocity is reduced near the centre, as illustrated in figures 7 and 8. According to (5.19), stagnation on the axis occurs when $C = 1$; and it can be seen that this case arises when the primary flow is sufficiently far supercritical, say when $\xi = \xi_s$. For $\xi > \xi_s$, the conjugate state is characterized by a central region of reversed flow.

By considerations precisely similar to those made in the previous example, the two propagation velocities of long travelling waves are found to be

$$c_+ = (\omega/\omega_c) + 1, \quad c_- = -(\omega/\omega_c) + 1. \tag{5.27}$$

Hence the characteristic number defined by (3.4) is identified as

$$N = \omega_c/\omega = \xi. \tag{5.28}$$

In particular we have $N = 1{\cdot}20^+/\omega$ for $R = \infty$. And as an obvious generalization of this result, when in an arbitrary system of units the axial velocity is W^*, the angular velocity of the core ω^*, and the core radius r^*, we have

$$N = 1{\cdot}20^+ \frac{W^*}{\omega^* r^*} \tag{5.29}$$

for a flow of infinite radial extent.

It is noteworthy that the present example allows a direct verification of the fundamental property of a conjugate flow, namely that it conserves the ψ-distributions of both circulation and total head given for the primary flow. In region I the assumed conditions give $I(\psi) = 2\omega^2\psi^2$ and $H(\psi) = H_0 + 2\omega^2\psi$. But $H = p/\rho + \frac{1}{2}(W^2 + V^2)$ can be calculated directly as a function of r by substituting (5.19) for W, using (A17) in the Appendix to find p, and putting $V^2 = 2I(\psi)/r^2$ with (5.18) substituted for ψ. After some straightforward reductions, the result is identifiable with the given form of $H(\psi)$ with (5.18) substituted for ψ. A corresponding check for region II is easily made since both H and I are constants (i.e. $V = \omega/r$ as in the primary flow). This convincing exercise is to be recommended if the general arguments of § 4 still leave doubt as to the possibility of distinct flows for which $H(\psi)$ and $I(\psi)$ are mutually conserved.

Another feature which is amendable to direct calculation in this example is the difference in flow-force between the two states A and B. The integral (4.4) for S may be evaluated explicitly both for the primary flow A, which is a very easy task, and also for the conjugate B when (5.18) is substituted for ψ and (5.19) for $W = \psi_y$ in region I. The calculation is straightforward though fairly lengthy, using standard results for various integrals involving Bessel functions, and it seems fair to omit the details here. The end-result is

$$S_B - S_A = \pi\rho\omega^2(\tfrac{3}{4} - \xi^2 + \tfrac{1}{4}\xi^4 + \log\xi), \tag{5.30}$$

which, rather surprisingly, does not involve R explicitly. The right-hand side of (5.30) is positive for $\xi > 1$; and in the light of our findings above, this result confirms the general theorem that the subcritical member of a conjugate flow pair has the greater flow-force.

Putting $\xi = 1 + \delta$ and expanding the terms binomially, we get from (5.30)

$$S_B - S_A = \pi\rho\omega^2(\tfrac{4}{3}\delta^3 + \tfrac{1}{5}\delta^5 - \tfrac{1}{6}\delta^6 + \ldots), \tag{5.31}$$

which accords with the general principle that the increase in flow-force associated with a supercritical-subcritical jump is of third order in the displacement produced (cf. the corresponding result (2.2) for open-channel flows).

6. Conclusion

The theory has provided a general interpretation of the essential factors in the vortex breakdown phenomenon, having demonstrated the relevant properties of swirling flows collectively in qualitative fashion rather than having analysed any particular flow model in detail—an alternative approach which would anyway present great difficulties if a reasonably accurate model for some actual flow were considered. Our explanation of dissipative breakdowns by reference to the energy-conserving conjugate flows is admittedly somewhat oblique; but it is logical, as the analogous explanation for hydraulic jumps clearly showed, and it has the great merit of dealing comprehensively with the widely various physical possibilities for the primary flow. To deal directly with the case of a dissipative vortex breakdown (i.e. to formulate a counterpart to Rayleigh's theory of the dissipative hydraulic jump), it would be necessary to introduce some additional hypothesis for the way in which the total-head loss due to turbulence is distributed over the stream-surfaces, and also a hypothesis regarding the diffusion of angular momentum by turbulent mixing. In consequence of these hypothetical factors, the comprehensiveness and precision of the present method of explanation would certainly be lost.

As regards practical applications, an important aspect of the theory is its status in relation to swirling flows whose structure varies continuously in the axial direction—e.g. leading-edge vortices which are well known to be susceptible to breakdown. Here the position is much the same as it is with the sample theory of normal shock waves, or the corresponding theory of hydraulic jumps; that is, after finding the properties of a 'discontinuity' between the two states of flow each with infinite length, one may usefully match the theory to local conditions in a varying flow, provided the scale of the variation is reasonably large in comparison with the length scale of the discontinuity.

Mr N. C. Lambourne has drawn my attention to several features which he observed in experiments at the National Physical Laboratory on breakdowns of leading-edge vortices, and which are not in direct accord with the present theory. The flow immediately following a breakdown generally appeared to be unsteady and not axisymmetric, in both respects differing from the case observed by Harvey (1962) in a 'vortex tube'. The outstanding observation was that the core of the original flow, marked with a smoke trace, was deformed into a spiral following the breakdown. It is felt that these features are in fact not inconsistent with present ideas, since they are probably accountable to lateral fluctuations superposed on a steady axisymmetric configuration. Thus, although the essential mechanism of vortex breakdown is explainable in terms of an axisymmetric model, there may be in practice considerable disturbances from this basic situation.

Appendix. Basic theory of axisymmetric swirling flows

(a) *The equation for the stream-function*

We take cylindrical co-ordinates (r, θ, x) with x along the axis of symmetry. In terms of orthogonal components respective to the co-ordinate directions, the velocity vector is denoted by $\mathbf{q} = (u, v, w)$ and the vorticity vector by $\boldsymbol{\omega} = (\xi, \eta, \zeta)$. The flow being steady, all velocity and vorticity components are functions of r and x only. The equation of continuity is therefore

$$u_r + r^{-1}u + w_x = 0, \tag{A 1}$$

which shows there to be a stream-function $\psi(r, x)$ such that

$$u = -r^{-1}\psi_x, \quad w = r^{-1}\psi_r. \tag{A 2}$$

Hence the resultant $\mathbf{q}_*$ of the velocity components u and w in a meridional plane is directed along the axisymmetric 'stream-surfaces' $\psi(r, x) = \text{const.}$, and the fluid particle paths are lines, generally spirals, on these surfaces. The magnitude of $\mathbf{q}_*$ is

$$q_* = -r^{-1}(\partial\psi/\partial n), \tag{A 3}$$

where n denotes the normal to the stream-surfaces.

Choosing to approach the dynamical problem in a somewhat indirect way that has the advantage of providing insight into the geometrical aspects, we take as three basic precepts the vector equation

$$\mathbf{q} \times \boldsymbol{\omega} = \nabla H \tag{A 4}$$

in which $H = p/\rho + \frac{1}{2}q^2$ (Milne-Thomson 1950, §3.45), Kelvin's circulation theorem, and its corollary that in a steady flow a vortex line must lie along a stream-surface. First we note that, when applied to a circuit around a particular stream-surface $\psi = \text{const.}$, Kelvin's theorem shows rv to be a constant; thus, in general,

$$rv = K(\psi), \tag{A 5}$$

where K is a function of ψ alone. Accordingly, $\mathbf{q}$ may be considered to be the resultant of the perpendicular components $v = K/r$ and q_* in a stream-surface (see figure A 1).

The component of vorticity in the θ-direction is given by (Milne-Thomson, p. 60)

$$\eta = u_x - w_r = -r^{-1}(\psi_{rr} - r^{-1}\psi_r + \psi_{xx}). \tag{A 6}$$

In consequence of the third basic result noted above, the resultant $\boldsymbol{\omega}_*$ of ξ and ζ must be directed along a stream-surface; and its magnitude ω_* is readily deduced by considering an infinitesimal surface normal to it drawn between consecutive stream-surfaces and two meridional planes subtending an angle $\delta\theta$. The area of this surface is $r\delta\theta\,.\,\delta n$, and the circulation round its boundary is $-\delta\theta\,.\,(\partial K/\partial n)\,\delta n$, which must equal ω_* times the area; thus

$$\omega_* = -r^{-1}(\partial K/\partial n) = q_* K'(\psi). \tag{A 7}$$

The total vorticity $\boldsymbol{\omega}$ is the resultant of the perpendicular components η and ω_* in a stream-surface (see figure A 1).

Now, (A 4) shows the gradient of H to be perpendicular to $\mathbf{q}$ and $\boldsymbol{\omega}$, which implies that it is normal to the stream-surfaces and hence that H is expressible as a function of ψ alone. We can rewrite (A 4) in the form

$$\partial H/\partial n = q\omega \sin\beta, \tag{A 8}$$

where β is the angle between $\mathbf{q}$ and $\boldsymbol{\omega}$ as shown in figure A 1; and by (A 3) this gives

$$rH'(\psi) = -(q/q_*)\,\omega\sin\beta = -\omega\sin\beta\sec\alpha, \tag{A 9}$$

where α is the other angle defined in figure A 1. But we have

$$\left.\begin{aligned} \omega\sin\beta &= \eta\cos\alpha - \omega_*\sin\alpha, \\ \text{and} \qquad \tan\alpha &= K(\psi)/rq_*. \end{aligned}\right\} \tag{A 10}$$

Hence, using also (A 6) and (A 7), and finally introducing the notation $I = \frac{1}{2}K^2$ so that $K(\psi)K'(\psi)$ may be written as $I'(\psi)$, we obtain directly from (A 9)

$$\psi_{rr} - r^{-1}\psi_r + \psi_{xx} = r^2 H'(\psi) - I'(\psi), \tag{A 11}$$

which is the required equation for $\psi(r, x)$. Alternative derivations proceeding rather more formally from the basic equations of motion have been given by Long (1953) and Squire (1956).

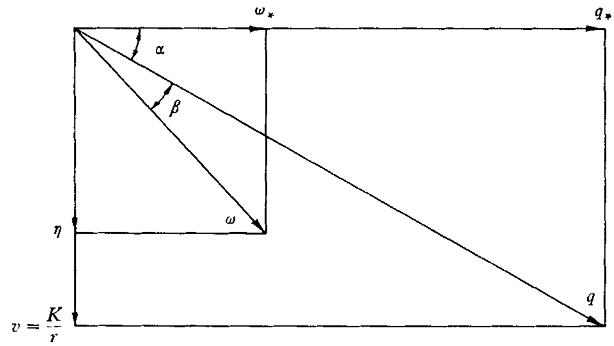

FIGURE A 1. Diagram of velocity and vorticity vectors in a stream-surface.

At this point it is convenient to introduce $y = \frac{1}{2}r^2$, the independent variable used in the main text. Using the fact that $\partial/\partial y \equiv r^{-1}(\partial/\partial r)$, we get from (A 11)

$$\psi_{yy} + (2y)^{-1}\psi_{xx} = H'(\psi) - (2y)^{-1}I'(\psi). \tag{A 12}$$

(b) *Perturbed cylindrical vortices*

We consider a primary flow with $u = 0$, $v = V(y)$, $w = W(y)$ and suppose a small steady disturbance to be superposed on it. The stream-function is accordingly expressed as

$$\psi = \Psi(y) + \epsilon\tilde{\psi}(y, x), \tag{A 13}$$

where Ψ is the stream-function for the primary flow and $\epsilon\tilde{\psi}$ is the perturbation. Note that Ψ is separately a solution of (A 12). Using this fact after substituting (A 13) into (A 12) and approximating the right-hand side to the first order in ϵ, we obtain

$$\tilde{\psi}_{yy} + (2y)^{-1}\tilde{\psi}_{xx} - \{H''(\Psi) - (2y)^{-1}I''(\Psi)\}\tilde{\psi} = 0. \tag{A 14}$$

In order to express the coefficient of $\tilde{\psi}$ in (A 14) explicitly as a function of y, it is first observed that for the primary flow

$$\left.\begin{aligned} W &= d\Psi/dy, \quad V^2 = I/y, \\ \text{and} \quad H &= p/\rho + \tfrac{1}{2}W^2 + \tfrac{1}{2}I/y. \end{aligned}\right\} \tag{A 15}$$

Hence
$$\frac{dH}{d\Psi} = \frac{p_y}{\rho W} + W_y + \frac{1}{W}\left(\frac{I_y}{2y} - \frac{I}{2y^2}\right). \tag{A 16}$$

But in a steady cylindrical flow the equation of radial equilibrium is

$$p_r = \rho V^2/r, \quad \text{or} \quad p_y = \tfrac{1}{2}\rho I/y^2. \tag{A 17}$$

This shows that the first and fourth terms cancel on the right-hand side of (A 16), so that

$$dH/d\Psi = W_y + \tfrac{1}{2}I_y/yW. \tag{A 18}$$

Again using the fact that $d/d\Psi \equiv W^{-1}(d/dy)$ to transform the second derivatives, it is now a simple matter to complete the reduction of (A 14). The result is

$$\tilde{\psi}_{yy} + \frac{1}{2y}\tilde{\psi}_{xx} - \left(\frac{W_{yy}}{W} - \frac{I_y}{2y^2W^2}\right)\tilde{\psi} = 0. \tag{A 19}$$

We can arbitrarily set $\psi = 0$ on the axis, and also $\Psi(0) = 0$. If the fluid is bounded by a cylindrical rigid surface at $y = a$, then clearly the total rate of flow is $Q = 2\pi\psi_{y=a}$. Assuming that Q is unchanged when the primary flow is perturbed, we therefore have

$$\tilde{\psi}(0,x) = 0, \quad \tilde{\psi}(a,x) = 0 \tag{A 20}$$

as boundary conditions on $\tilde{\psi}$.

(c) *Wave resistance*

We proceed to investigate the change in flow-force caused by perturbing a cylindrical flow in the manner assumed above—that is, under the conditions that Q and the distributions of H and I over the stream-surfaces are unchanged. The flow-force S is the integral over the cross-section of the quantity

$$p + \rho u^2 = \rho\{H(\psi) + \tfrac{1}{2}u^2 - \tfrac{1}{2}w^2 - \tfrac{1}{2}I(\psi)/y\},$$

and we denote its values before and after perturbation by S_1 and S_2, respectively Substituting (A 13) into the integrand for S_2 and approximating to the second order in ϵ, we obtain

$$S_2 - S_1 = 2\pi\rho\int_0^a \left[\epsilon\left\{H'(\psi) - \frac{1}{2y}I'(\Psi)\right\}\tilde{\psi} + \epsilon W\tilde{\psi}_y + \tfrac{1}{2}\epsilon^2\left\{H''(\Psi) - \frac{1}{2y}I''(\Psi)\right\}\tilde{\psi}^2 + \tfrac{1}{2}\epsilon^2\tilde{\psi}_y^2 - \frac{1}{4y}\epsilon^2\tilde{\psi}_x^2\right]dy. \tag{A 21}$$

Since $H'(\Psi) - I'(\Psi)/2y = W_y$ (see equation (A 18)), the two first-order terms in the integrand reduce to the single y-derivative $(W\tilde{\psi})_y$, and so their integral vanishes in consequence of the boundary conditions (A 20). A substitution for the term in $\tilde{\psi}^2$ can be made from (A 14), and two terms then appearing in the integrand reduce to $\frac{1}{2}(\tilde{\psi}\tilde{\psi}_y)_y$ so that their integral also vanishes. Hence we are led to

$$S_2 - S_1 = 2\pi\rho\,\epsilon^2\int_0^a \frac{1}{4y}(\tilde{\psi}\tilde{\psi}_{xx} - \psi_x^2)\,dy. \tag{A 22}$$

Consider now a standing-wave disturbance with

$$\tilde{\psi} = \phi(y)\sin(\alpha x + \nu). \tag{A 23}$$

By (A 19), ϕ must be a solution of

$$\frac{d^2\phi}{dy^2} - \left(\frac{\alpha^2}{2y} + \frac{W_{yy}}{W} - \frac{I_y}{2y^2W^2}\right)\phi = 0. \tag{A 24}$$

Substitution of (A 23) into (A 22) shows at once that the *reduction* in S due to wave formation is

$$\Delta S = S_1 - S_2 = \tfrac{1}{2}\pi\rho\epsilon^2\alpha^2 \int_0^a \frac{\phi^2}{y}\, dy, \tag{A 25}$$

which is always positive and is independent of x, as is to be expected. This is the quantity described as the wave resistance in the main text.

Note that when $\tilde{\psi}$ has an exponential dependence on x the integral (A 22) vanishes, which proves S to be unchanged by the disturbance. This property is obviously to be expected on physical grounds, since far in front of an axisymmetrical obstacle fixed in a cylindrical vortex the disturbance created will diminish exponentially with distance, yet in the absence of any external force acting in the axial direction the flow-force must be constant everywhere ahead of the obstacle.

(d) The group velocity of standing waves

To establish this property of standing waves, we need to consider their relation to the class of *travelling* waves for which the stream-function perturbation has the form

$$\phi(y) \sin[\alpha(x - ct) + \nu], \tag{A 26}$$

where c is phase velocity and t time. Since this wave motion will appear steady in a frame of reference in which the primary velocity components are V and $W - c$, therefore ϕ must satisfy an equation of the form (A 24) with $W - c$ in place of W. The boundary conditions are $\phi(0) = \phi(a) = 0$ as before.

Here we are particularly interested in values of c close to zero, corresponding to which α^2 must necessarily be close to a positive eigenvalue, say α_0^2, of the Sturm–Liouville system comprising the standing-wave equation (A 24) and the boundary conditions (see § 3). It is assumed, of course, that the flow is subcritical so that a positive α_0^2 does exist.† We shall also assume that $W > 0$ throughout the interval $(0, a)$. Our primary aim is to show that the variation of α^2 with c imposed by the eigenvalue problem with c a free parameter is such that the group velocity defined as

$$d(\alpha c)/d\alpha = c + 2\alpha^2(dc/d\alpha^2) \tag{A 27}$$

is generally positive (i.e. in the flow direction) for $c \to 0$.

Identifying the standing-wave solution by a zero suffix we write $\phi_0 = Wf_0$. This substitution in (A 24) leads to

$$\frac{d}{dy}\left(F_0 \frac{df_0}{dy}\right) - G_0 f_0 = 0, \tag{A 28}$$

with
$$F_0 = W^2, \quad G_0 = \tfrac{1}{2}\alpha_0^2 W^2/y - \tfrac{1}{2}I_y/y^2. \tag{A 29}$$

† Note incidentally that at least if $I_y > 0$ there is always, for any assigned value of α^2, an admissible value of c which is greater than the maximum of W. This fact appears on consideration that the coefficient of ϕ in the travelling-wave equation is continuously adjustable between $-\tfrac{1}{2}\alpha^2/y$ and $+\infty$ by such a choice of c. In contrast with the waves thus indicated, which propagate with the flow, the waves which mainly concern us here are, of course, ones propagating against the flow. (For a discussion of rather similar ideas relating to water waves, see Brooke Benjamin (1962, § 5).)

Writing the corresponding travelling-wave solution as $\phi_1 = (W-c)f_1$, we have that f_1 satisfies an equation like (A 28) but with modified coefficients F_1, G_1 which are given by replacing W with $W-c$ in the definitions (A 29).

Suppose now that the phase velocity is made an infinitesimal δc, which in the first place may be taken as positive. For the eigenvalue of the travelling-wave system we may write $\alpha_1^2 = \alpha_0^2 + \delta\alpha^2$, where $\delta\alpha^2$ is the variation dependent on δc. Since $W > 0$, we have that $F_0 > F_1$ throughout $(0, a)$; and if $d\alpha^2/dc$ were either *zero or negative* we would necessarily have also that $G_0 > G_1$ throughout $(0, a)$. But according to Picone's generalization of Sturm's fundamental comparison theorem (Ince 1926, §10.31), these two conditions imply that f_1 would oscillate more rapidly that f_0 in this interval; in other words, whereas zeros of f_0 lie at the end-points of the interval, corresponding zeros of f_1 would span an interval less than the original by an infinitesimal amount of the order of δc. Yet f_1 cannot have successive zeros which are not finitely spaced (Ince, §10.2). Thus, since f_1 as well as f_0 has to vanish at both end-points, the case under test is proved impossible. Hence $d\alpha^2/dc$ must be positive; and clearly one would arrive at the same conclusion taking $\delta c < 0$.

This establishes that the group velocity defined by (A 27) is positive in the limit $c \to 0$. The important physical significance of this result has been explained in the main text.

A similar argument can readily be applied to certain cases of travelling waves with α^2 different by a finite amount from an eigenvalue α_0^2 for standing waves, and serves in particular to verify another important property which has been assumed, namely that waves of extreme length (i.e. with $\alpha^2 \to 0$) can propagate upstream under subcritical conditions. Consider the largest eigenvalue, say α_m^2, possible for a standing wave—which may of course be the only one if the flow is just marginally subcritical (see §3). The corresponding solution f_0 will oscillate exactly once over the interval $(0, a)$. If we were to put $\alpha^2 < \alpha_m^2$ in (A 28) leaving W unchanged, we would certainly find a solution which oscillates *within* $(0, a)$ since, while the F coefficient is unchanged, the G coefficient is made everywhere more negative than G_0 in the standing-wave equation, thus increasing the rate of oscillation of the solution (cf. Ince, §10.3). But the rate of oscillation can now be decreased to an arbitrary extent by replacing W^2 with $(W-c)^2 > W^2$, both because the F coefficient is made larger and the G coefficient less negative. Hence there is always a choice of a negative value for c which will prolong the oscillation up to the extent of $(0, a)$. There will in fact be more than one choice of c to satisfy the boundary conditions if the solution for $\alpha^2 < \alpha_m^2$ and W unchanged oscillates more than once within $(0, a)$; but clearly the largest of the critical values of $-c$ is that which prolongs the first oscillation up to $(0, a)$. It is thus demonstrated that waves with $\alpha^2 < \alpha_m^2$ exist which can propagate upstream. In particular, the largest propagation velocity $-c$ is obtained for $\alpha^2 \to 0$ and when the respective solution f_1 makes just one oscillation over $(0, a)$.

REFERENCES

BENJAMIN, T. BROOKE 1956 On the flow in channels when rigid obstacles are placed in the stream. *J. Fluid Mech.* **1**, 227.

BENJAMIN, T. BROOKE 1962 The solitary wave on a stream with an arbitrary distribution of vorticity. *J. Fluid Mech.* **12**, 97.

BENJAMIN, T. BROOKE & LIGHTHILL, M. J. 1954 On cnoidal waves and bores. *Proc. Roy. Soc.* A, **224**, 448.

BOLZA, O. 1961 *Lectures on the Calculus of Variations.* New York: Dover.

BURNS, J. C. 1953 Long waves in running water. *Proc. Camb. Phil. Soc.* **48**, 695.

ELLE, B. J. 1960 On the breakdown at high incidences of the leading-edge vortices on delta wings. *J. Roy. Aero. Soc.* **64**, 596.

FOX, C. 1950 *Calculus of Variations.* Oxford University Press.

HALL, M. G. 1961 A theory for the core of a leading-edge vortex. *J. Fluid Mech.* **11**, 209.

HARVEY, J. K. 1960 Analysis of the vortex breakdown phenomenon, Part II. *Aero. Dept., Imperial Coll., Rep.* no. 103.

HARVEY, J. K. 1962 Some observations of the vortex breakdown phenomenon. *J. Fluid Mech.* **14**, 585.

INCE, E. L. 1926 *Ordinary Differential Equations.* London: Longmans. (Dover edition 1956.)

JONES, J. P. 1960 The breakdown of vortices in separated flow. *Dept. Aero. Astro., Univ. of Southampton, Rep.* no. 140.

LAMB, H. 1932 *Hydrodynamics*, 6th ed. Cambridge University Press.

LONG, R. R. 1953 Steady motion around a symmetrical obstacle moving along the axis of a rotating fluid. *J. Met.* **10**, 197.

LUDWIEG, H. 1961 Contribution to the explanation of the instability of vortex cores above lifting delta wings. *Aero. Versuchsanstalt, Gottingen, Rep.* AVA/61 A01.

MILNE-THOMSON, L. M. 1950 *Theoretical Hydrodynamics*, 2nd ed. New York: Macmillan.

SQUIRE, H. B. 1956 *Rotating fluids.* Art. in *Surveys in Mechanics* (Ed. Batchelor and Davies). Cambridge University Press.

SQUIRE, H. B. 1960 Analysis of the vortex breakdown phenomenon, Part I. *Aero. Dept., Imperial Coll., Rep.* no. 102.

WERLÉ, H. 1960 Sur l'eclatement des tourbillons d'apex d'une aile delta aux faibles vitesses. *Res. Aero., Paris*, no. 74, 23.