first one, which immediately gave the tint of the A constituent in his formulas. And as it varied with the thickness, A ran through the reflection tints in Newton's scale. This experiment, in which one uses a second polarizing reflection, corresponds to setting α equal to 0° when, instead, one uses an analyzing crystal: for in the latter case Biot's formulas imply that the extraordinary image in the crystal will consist solely of the A beam. With α equal to 0°, then, one may say that the extraordinary image should correspond to Newton's reflection rings as the thickness of the lamina changes.

If, therefore, the extraordinary image corresponds to Newton's scale of *reflection* tints, then by energy conservation the ordinary image must correspond to Newton's scale of transmission tints. Fresnel had earlier shown that to explain Newton's reflection rings one must add a half-wavelength path difference. By analogy, for α zero one must add a half-wavelength to the actual path difference between the constituents of the extraordinary beam. If one changes α by 90°, then by the same reasoning this addition must be made to the path difference between the constituents of the ordinary beam.

Fresnel went a bit further. What, he apparently asked himself, distinguishes the case of α equal to 0° from that of α equal to 90°, where one must switch the additional half-wavelength from the extraordinary image to the ordinary one? In the first case, he noted, α is smaller than i, whereas in the other it is larger. That is, in the first case the analyzer's principal section lies between the original plane of polarization and the optic axis of the lamina; in the second case the lamina's axis lies between the analyzer's principal section and the plane of polarization. Fresnel then assumed—quite without justification in Biot's observations, since Biot never altered α at all—that *in general* one must add a half-wavelength to the path difference between the constituents of the extraordinary beam when α is less than i, and to the difference for the ordinary beam when α is greater than i.

Fresnel insisted on generalizing from Biot's observations, because he was reaching for some sort of understanding of how polarization operates in conjunction with the principle of interference. Accordingly he did not stop with this claim, but noted that it is equivalent to another one, which he expressed in the following words: "The image whose tint corresponds exactly to the thickness of the crystal lamina is that whose two constituent beams have each undergone two opposite motions in their plane of polarization; while in the two beams that produce the complementary tint, the plane of polarization, on the contrary, always deviates in the same direction from its original position" (Fresnel 1816c, 402). The extra phase difference, then, arises whenever the planes of polarization of the constituent beams twice increase the angle between them—and so come to lie again along the same line, after having deviated from one another through 180°. Fresnel later remarked, we have already seen, that he and Ampère had early realized that this fact could be understood if the waves were entirely transverse, since then an angle of 180° between the planes of polarization equates to direct opposition between the directions of oscillation.

APPENDIX 17 Decomposing Reflected and Refracted Light

17.1 Energy and Momentum Conservation for Waves

We have already seen how Fresnel used energy principles to obtain results that could not be found directly. His first demonstration that the images produced by chromatic polarization must have complementary colors was based on energy conservation, and he often recurred to the principle when he could not draw a result directly from calculation.[16] And his first attempt to derive reflection formulas, in mid-1819, employed energy conservation as a check against results obtained from another conservation principle—that of momentum.

In the *Oeuvres* (1:649–53) Verdet reproduces a note, itself undated, contained in a folder marked 12 July 1819. Given its contents, the note was almost certainly written about this time. In it Fresnel attempts to derive reflection formulas by deploying momentum conservation and then checking the results against conservation of energy. Throughout the analysis he assumes that the polarized wave consists entirely of a *longitudinal* oscillation—which therefore seems to show that he had not by this date developed the idea that even unpolarized light can be completely asymmetric (see fig. A17.1).

In Fresnel's figure GA and HC are two rays from the incident wave front, which we assume to be plane. AM and CN are their respective reflections and AK, CL the refractions. Fresnel's goal was to obtain expressions for the reflected (v) and refracted (u) amplitudes in terms of the incident amplitude (V), the angle of incidence (i), and the index of refraction (n). Obviously, to do so two equations are necessary, and Fresnel generated them by adapting the requirement of momentum conservation to waves.

The problem was how to choose the "masses" to use in computing the momenta. Fresnel assumed that the momentum a given mass in the incident beam has at a given time will be transmitted in a time δt, during which the incident front traverses this mass, to the masses in the reflected and refracted beams, which are themselves traversed by the reflected and refracted fronts during δt. In the figure, during the time that point E moves to C on the interface, the volume AEC is traversed by the front AE. During this same time the volumes ADC (equal to AEC) in the reflected, and ABC, in the refracted, beams are traversed by the fronts DC and BC.

If the masses of AEC and ADC are m and the mass of ABC is μ, then momentum conservation (or what Fresnel termed the "conservation of the motion of the center of

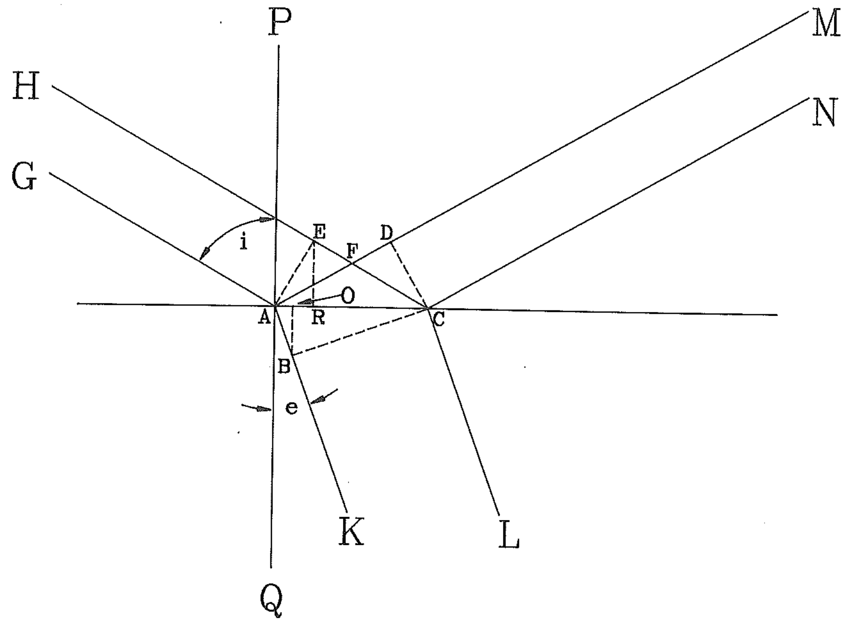

FIG. A17.1 Fresnel's diagram for energy and momentum conservation in reflection.

gravity") yields two equations, one for the vertical, the other for the horizontal component of the momentum:

$$Vm\sin(i) = u\mu\sin(e) + vm\sin(i)$$

$$Vm\cos(i) = u\mu\cos(e) - vm\cos(i)$$

The angle of incidence, i, is given, and the angle e of refraction follows from Snel's law. Consequently only the masses m and μ must be eliminated. Fresnel assumed that the density of the refracting medium is proportional to the square of its index of refraction.[17] Then geometry yields $n\cos(e)/\cos(i)$ for the mass ratio μ/m. Substituting in the momentum equations and solving yields in the end:

$$v = V\frac{\sin(i - e)}{\sin(i + e)}$$

$$u = V\frac{\cos(i)}{r\cos(e)}\frac{\sin(2i)}{\sin(i + e)}$$

Note that these expressions represent the amplitudes of oscillations that are normal to the front.

Fresnel next checked his results against energy conservation. As he understood it the latter required the energy (vis viva) in AEC to equal the sum of the energies in ADC and ABC, or:

$$mV^2 = mv^2 + \mu u^2$$

Combining our three equations with Snel's law yields in the end:

$$\cos^2(e) = \cos^2(i)$$

which, unless the index of refraction is unity, can only be satisfied when i is equal to e, that is, when the angle of incidence vanishes. Consequently Fresnel's equations are consistent with energy conservation only at zero incidence.

Either, Fresnel concluded, Snel's law itself does not hold generally or else the waves may "also" contain transverse oscillations, in which case the entire analysis must be redone:

> With the help of transverse motions one could satisfy both the general principle of the conservation of the center of gravity and that of the conservation of living force, which must be satisfied in all the vibrations of elastic fluids, and one might be able, in thus determining the transverse motions of waves, to define that singular modification of light to which we have given the name *polarization.* (Fresnel, *Oeuvres,* 1:652)

That is, by working backward from momentum and energy conservation one might, Fresnel seems to reason, be able closely to specify the oscillatory structure of the wave. That had become a moot point by 1821, but Fresnel preserved the essential features of this early derivation of the sine law ratio in his first analysis for polarized light.

17.2 Decomposition

In his new analysis, which was based directly on the requirement that the oscillation must occur entirely in the wave front, Fresnel continued (in effect) to rely on momentum conservation, but he had also to use energy principles because the oscillation is no longer parallel to the ray. If the oscillation is along the ray, then momentum conservation, as we have seen, provides the two necessary equations. But now the oscillation must be normal to the ray. Suppose it to occur entirely in a plane parallel to the reflecting interface. Then momentum conservation will not apparently involve the angles i and e at all, because here we have no component in the plane of incidence itself.

Fresnel had accordingly to modify his previous analysis. In 1821 he provided the first of two ways to do so, this one being (in his eyes and in those of later wave theorists) the less satisfying of the two.[18] The essence of both procedures requires decomposing the incident oscillation into two components. One lies in the plane of incidence; the other is perpendicular to it. At first Fresnel was able to solve only the case in which the oscillation occurs in the normal to the plane of incidence, whereupon it is parallel to the interface. He again used momentum conservation, but he had to alter his previous approach, which had applied to longitudinal oscillations.

In this case of an oscillation parallel to the interface we have, Fresnel assumed, to deal with the elastic impact between two bodies along the line joining them. The effect of this assumption is to combine energy with momentum conservation (though Fresnel would not have thought about it in quite that way). In the previous case of an oscillation along the ray we had to consider independently the two components of the impact, since the direction of the oscillation alters as a result of the interaction. Here the direction of the oscillation remains the same; only the direction of propagation changes. Suppose that the masses and velocities are represented as above: that is, we treat the portions of the incident, reflected, and refracted beams that Fresnel had previously delineated as colliding "masses." Then the usual formulas for an elastic collision require that, after the impact is complete, the velocity (amplitude) ratios must be:

$$\frac{v}{V} = \frac{m - \mu}{m + \mu} \quad \text{and} \quad \frac{u}{V} - \frac{2m}{m + \mu}$$

Since the mass ratio remains $n(\cos(e)/\cos(i))$ we have for the optical intensity ratio (the square of the amplitude ratio):

Fresnel's Sine Law
oscillation normal to the plane of incidence

$$\frac{\text{reflected intensity}}{\text{incident intensity}} = \left[\frac{\tan(i) - \tan(e)}{\tan(i) + \tan(e)}\right]^2 = \left[\frac{\sin(i - e)}{\sin(i + e)}\right]^2$$

This is precisely the same as the ratio Fresnel had previously found for longitudinal oscillations.

To this point Fresnel could not compare his formula with experiment because he was not as yet able to deduce the corresponding formula for an oscillation that occurs in the plane of incidence. He remarked that, from two measurements Arago had made of the deviation of the plane of polarization by reflection, one could use the new formula to deduce for both instances the amplitude of the plane-parallel component of the reflection, and thence the total intensity of the reflected light. However, the only photometric measures available were old ones generated by Pierre Bouguer, and they were too inaccurate for quantitative comparisons.

Within a few days at most after finishing the "Calcul des teintes" paper, in which he detailed his new hypothesis that light is always asymmetric, Fresnel remarked in a postscript that he had "by a mechanical solution" obtained the formula for an oscillation that takes place in the plane of incidence, finding it to be:

Fresnel's Tangent Law
oscillation in the plane of incidence

$$\frac{\text{reflected intensity}}{\text{incident intensity}} = \left[\frac{\sin(2i) - \sin(2e)}{\sin(2i) + \sin(2e)}\right]^2$$

(Fresnel later reduced the ratio to the form $[\tan(i - e)/\tan(i + e)]^2$. Note that in this

analysis he did not specify the sign of the amplitude ratio itself, since he gave only its square.)

The first formula, for light oscillating parallel to the interface, must hold for polarization in the plane of incidence, because the reflected intensity cannot vanish at any angle (excepting the trivial case of normal incidence). Consequently, on the basis of this analysis the oscillation takes place along the normal to the plane of polarization. The second ratio, for light oscillating in the plane of incidence, can vanish. From the equivalent tangent formula we at once see that this will occur when the angles of incidence and refraction are complementary (thereby satisfying Brewster's law that the light will be completely polarized when the tangent of the incidence is equal to the index of refraction—a point Fresnel noted in his table comparing theory with experiment).

Fresnel did not describe his "mechanical solution" here, but it was no doubt the same as the one he provided a year and half later in his last, and most detailed, analysis of reflection (Fresnel 1823a). The new method applies as well to the sine law as to the tangent law and makes no use at all of elastic collisions. Instead it combines energy conservation *with a constraint on the continuity of the medium at the interface* and continues to assume, as before, that the density ratio varies with the square of the refractive index. In other words, Fresnel for the first time used a true boundary condition. Previously he had not actually considered what happens to the oscillation at the interface; he had assumed only that, whatever happens, energy and momentum are conserved. To do this he had used Snel's law to determine the cross sections of the incident, reflected, and refracted beams. He had then, in effect, pressed the energy and momentum ahead of a front in the incident beam into the reflected and refracted beams. He continues to deploy this method here, but he now combines it with a constraint on the actual oscillation at the boundary.

Thus Fresnel first used energy conservation together with his expression for the density ratio to obtain an equation that connects the amplitudes with the angles of incidence and refraction. The analysis is the same as his earlier one for the case of a longitudinal oscillation, since the only things to be determined are the corresponding volumes in the reflected and the refracted fronts that, in a given time δt, receive the energy from a given volume traversed during δt by the incident front. Consequently the energy equation is, for unit incident amplitude:

Fresnel's Energy Equation

$$\sin(e)\sin(i)(1 - v^2) = \sin(i)\cos(e)u^2$$

For a boundary condition Fresnel assumed that the components of the oscillation that are *parallel* to the interface must be continuous across it. That is, the medium does not support slip. Since we take the incident amplitude as one, this gives us, for the case of an oscillation that is itself parallel to the interface:

$$1 + v = u$$

Combined with the energy equation, this condition then yields the sine law by easy manipulation, including now the sign of the amplitude ratio proper:

$$v = -\frac{\sin(i - e)}{\sin(i + e)}$$

The boundary condition also suffices to solve the case for an oscillation that is parallel to the plane of incidence, for here we need only take the oscillation's component parallel to the interface, which gives as boundary condition:

$$\cos(i) + v\cos(i) = u\cos(e)$$

With the energy equation we obtain the tangent law in the form:

$$v = -\frac{\sin(i)\cos(i) - \sin(e)\cos(e)}{\sin(i)\cos(i) + \sin(e)\cos(e)}$$

Many readers will at once recognize these two laws as the very ones that are today obtained from electromagnetic principles (a deduction first explicitly carried out by H. A. Lorentz in 1875). One might be surprised by this, since the principles involved seem so vastly different. However, the deductions are rather closely related. Fresnel's energy equation expresses what, in electromagnetic theory, one deduces from the Poynting equation for the flux of energy in the field. And Fresnel's boundary condition corresponds to the electromagnetic requirement that the tangential component of the electric field must be continuous at the boundary. These two equations—Poynting's and the continuity of the tangential electric field—suffice to deduce the Fresnel amplitude ratios.

These two laws were extremely important historically because, as we shall see, anyone who accepted them both had also accepted, and understood, the deepest recesses of the wave theory. Indeed, for many years they remained *individually* quite difficult for many of Fresnel's contemporaries to grasp. They could not in any case be tested in themselves, since accurate photometric techniques were not available, so that until the late 1820s even those (like Arago himself) who were otherwise sympathetic to the wave theory tended to ignore them. What could be examined, and what could be understood by most everyone, however, was the angle of deviation for the incident direction of polarization that the two expressions together entail.

Suppose the incident wave is polarized at an angle α to the plane of incidence—which means, on Fresnel's analysis, that the oscillation occurs at an angle $90° - \alpha$ to it. Assume that the amplitude of the incident wave is unity, which therefore has components $\sin(\alpha)$, $\cos(\alpha)$, respectively, in and perpendicular to the plane of incidence. The tangent of the angle of polarization, α_{refl}, of the reflected wave will therefore be:

$$\tan(\alpha_{refl}) = \cot(\alpha)\frac{\cos(i - e)}{\cos(i + e)}$$

This formula could be tested directly (indeed Fresnel had, in effect, already done so for α equal to 45°), and by 1830 it had come into common use even among selectionists like Brewster.[19] But it was often taken—like Fresnel's diffraction formulas by Poisson—as a purely empirical rule without a proper theoretical basis.

17.3 Phase Shifts Where the Equations Fail

Fresnel's mastery over the process of composition and decomposition appears to its best advantage in his new interpretation of total reflection, an intepretation that builds on his previous understanding. If the index n is less than one (the incidence then being internal), then beyond an incidence of arcsine(n), the value of $\cos(e)$, and with it the Fresnel ratios, must become complex.[20] This is also the incidence beyond which total internal reflection occurs—all the light is reflected. This seems to have provided Fresnel with a clue for interpreting the meaning of complex amplitude ratios, and his work here was of great importance during the 1830s, when the wave theory was developed in great analytical detail.

The amplitude ratios must have become complex, Fresnel thought, because a condition that had gone into deriving them had been violated. That condition, he assumed, was the assumption that the phase is continuous at the interface (Fresnel 1823h, 783). Suppose instead that when the light strikes its coherence is momentarily interrupted. This will yield a reflection that is shifted in phase from the incident wave. Can we, Fresnel in effect queried, interpret the complex ratios in a way that would be consistent with such a phase shift? And if so, what must the new phase be?

To understand Fresnel's reasoning, first rewrite the ratio for light polarized perpendicular to the plane of incidence in a way that separates the real from the imaginary part:

$$v = -\frac{\sin(i - e)}{\sin(i + e)} = \frac{1 + n'^2 - 2n'^2\sin^2(i)}{n'^2 - 1} - \frac{2n'\cos(i)\sqrt{n'^2\sin^2(i) - 1}}{n'^2 - 1}\sqrt{-1}$$

(defining n' as the reciprocal of the index n from the more to the less refracting medium). Fresnel noticed that the sum of the squares of the real part (R^2) with the real factor (I^2) of the imaginary part is equal to one. That very likely suggested to him the following interpretation.

The light must be totally reflected, and it must undergo a phase shift φ_1. Since the reflected wave has some new phase, it can, by Fresnel's usual quarter-wave decomposition, be separated into two parts that differ in phase from one another by 90°. One of the parts has the same phase as the incident wave, and the amplitudes of the parts must be the cosine and sine of the phase of their resultant. Now here we have two terms, R and I, whose squares sum to one. Suppose that each of these represents the amplitude of one of the parts of the usual quarter-wave decomposition. Then the light will be completely reflected, and it will be shifted in phase by an angle φ whose tangent is R/I if we assume that R is the part with the same phase as the incident wave.[21]

One can do precisely the same thing for the component of the incident light that is polarized in the plane of incidence, computing for it a phase shift φ_2. Then, Fresnel concluded, the reflection will consist of two orthogonal components that differ in phase from one another by the quantity $\varphi_1 - \varphi_2$:

$$\cos(\varphi_1 - \varphi_2) = \frac{2n'^2\sin^4(i) - (n'^2 + 1)\sin^2(i) + 1}{(n'^2 + 1)\sin^2(i) - 1}$$

Furthermore, the amplitudes of the orthogonal components are each equal to one.[22] This result, Fresnel demonstrated in some detail, fully explains the depolarizing behavior that he had observed much earlier in totally reflecting light two times within his rhomb. In terms of his new vocabulary, this light has been *circularly polarized.*

APPENDIX 18 Arago's Critique of Biot on Chromatic Polarization

Set up a mirror experiment, but look only at the images formed by rays that reflect from each mirror near polarizing incidence. Then the fringe pattern is created by light that is polarized in the plane of incidence, as attested by an analyzer. Now, Arago continues, take two gypsum laminae and place each in the path of one of the two beams that interfere to produce the fringe pattern. Orient one of the laminae so that its axis lies at 45° to the plane of incidence; orient the other lamina's axis at right angles to this (see figs. A18.1 and A18.2).

If Fresnel's theory is correct, each thin lamina will behave precisely like a thick crystal and so will generate from the polarized beam that strikes it two others: one polarized for each lamina in its principal section, the other polarized in the section's normal. By construction the ordinary beam from the first lamina will have the same polarization as the extraordinary beam from the second lamina, and vice versa. Thus, in figure A18.2 the ordinary beam KMF from lamina A will interfere with the extraordinary beam NOF from lamina B, while the extraordinary beam KLF from A will interfere with the ordinary beam NPF from B. But the path difference between KLF and NPF is different from the path difference between KMF and NOF, so that the fringe patterns will not be coincident. There must accordingly be two patterns, one of which should be polarized at 45° to the plane of incidence, whereas the other should be polarized in the plane normal to the first. This, Arago remarks, is "perfectly confirmed" by experiment.

But if so, then Biot's theory must be wrong. For according to Biot, the light that emerges from both laminae is polarized in part in, and in part along the normal to, the plane of incidence (since here $2i$ is 90°). Consequently the fringes should be polarized in these two planes. Since they are not, Biot is wrong.

The gist of Arago's criticism is that one cannot possibly assume, as Biot did, that each of the two beams consists of light polarized in and perpendicular to the plane of incidence—because the fringe pattern shows completely different polarizations than this would require. This is indeed a powerful criticism, and Biot never answered it (but see my remarks in sec. 9.1).