

The Causes of Color

They are diverse, but they all stem from the same root: It is the electrons in matter, through their varied responses to different wavelengths of light, that make the world a many-colored place

by Kurt Nassau

What makes the ruby red? Why is the emerald green? On the most superficial level these questions can be given simple answers. When white light passes through a ruby, it emerges with a disproportionate share of longer wavelengths, which the eye recognizes as red. Light passing through an emerald acquires a different distribution of wavelengths, which are perceived as green. This explanation of color is correct as far as it goes, but it is hardly satisfying. What is missing is some understanding of how matter alters the composition of the light it transmits or reflects. Ruby and emerald both derive their color from the same impurity element: Why then do they differ so dramatically in color? What gives rise to the fine gradations in spectral emphasis that constitute the colors of materials?

It turns out that the ultimate causes of color are remarkably diverse. An informal classification I shall adopt here has some 14 categories of causes, and some of the categories embrace several related phenomena. With one exception, however, the mechanisms have an element in common: the colors come about through the interaction of light waves with electrons. Such interactions have been a central preoccupation of physics in the 20th century, and so it is no surprise that explanations of color invoke a number of fundamental physical theories. Indeed, color is a visible (and even conspicuous) manifestation of some of the subtle effects that determine the structure of matter.

The Energy Ladder

The perception of color is a subjective experience, in which physiological and psychological factors have an important part; these matters will not be taken up in detail here. It seems reasonable to assume, however, that perceived color is merely the eye's measure and the brain's interpretation of the dominant wavelength or frequency or energy of a light wave. The meaning of this assumption is clear in the case of monochromatic light, which has a single, well-defined

wavelength. The interpretation of light that is a mixture of many wavelengths is more complicated, but it is still the relative contributions of the various wavelengths that determine the color.

Wavelength, frequency and energy are alternative means of characterizing a light wave. Energy is directly proportional to frequency; both energy and frequency are inversely proportional to wavelength. In other words, high frequencies and high energies correspond to short wavelengths, as at the violet end of the visible spectrum. A common unit of measure for light wavelengths is the nanometer, which is equal to a billionth of a meter. The energy of light is conveniently measured in electron volts, one electron volt being the energy gained by an electron when it accelerates through a potential difference of one volt. In terms of wavelength human vision extends from about 700 nanometers, where red light grades into infrared radiation, down to about 400 nanometers, at the boundary between violet light and ultraviolet radiation. The same range in energy units runs from 1.77 electron volts to 3.1 electron volts.

An important constraint on all interactions of electromagnetic radiation with matter is the quantum-mechanical rule that says atoms can have only certain discrete states, each with a precisely defined energy; intermediate energies are forbidden. Each atom has a lowest-possible energy, called the ground state, and a range of excited states of higher energy. The allowed energy states can be likened to the rungs of a ladder, although their spacing is highly irregular. Light or other radiation can be absorbed only if it carries precisely the right amount of energy to promote an atom from one rung to a higher rung. Similarly, when an atom falls from an excited state to a lower-lying one, it must emit radiation that will carry off the difference in energy between the two levels. The energy appears as a photon, or quantum of light, whose frequency and wavelength are determined by the energy difference.

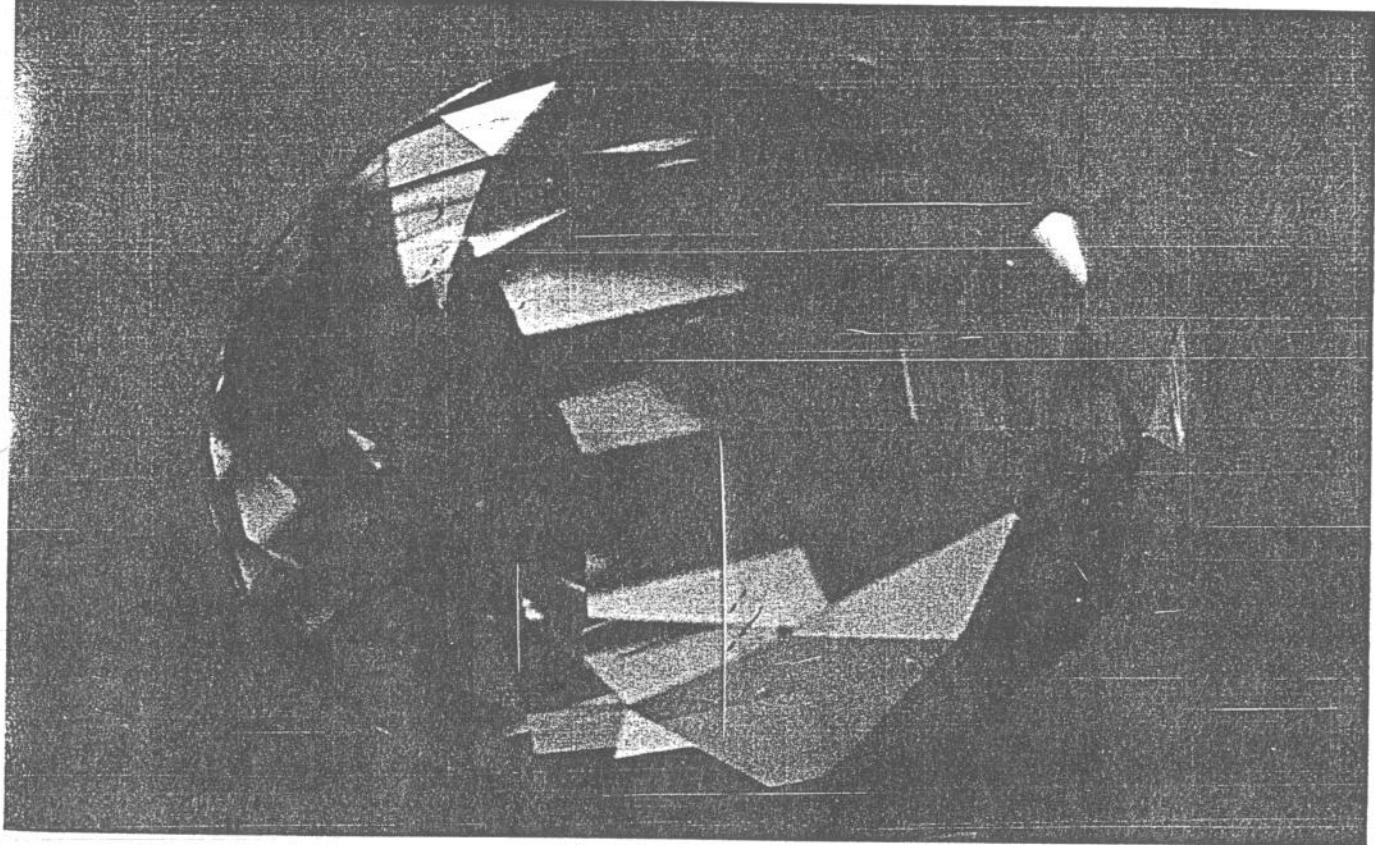
The states that are of the greatest in-

terest in the analysis of color represent various possible energy levels of electrons. In atoms, ions and molecules each electron must occupy an orbital, which describes a particular geometric distribution of the electron's charge around the atomic nucleus. The orbitals in turn are organized in shells. A further constraint on the possible states of the atom is that each rung on the energy ladder can be occupied by only a limited number of electrons. In general, when proceeding from the smallest atoms to the largest ones, electrons are added in sequence from the bottom rung up. Two electrons fill the first shell; each of the next two shells holds eight electrons. The electrons in any filled or closed shell form pairs, and they have a notably stable configuration.

A comparatively large quantity of energy is needed to promote one of the paired electrons from a closed shell to the next vacant position on the ladder. The energy required for such a transition can usually be supplied only by radiation in the ultraviolet or even in the X-ray region of the spectrum; as a result closed shells have no direct influence on the colors of materials. Instead color usually results from transitions of unpaired electrons, which are most often the outermost ones. They are the valence electrons, the ones that participate in chemical bonds.

Atomic Transitions

Consider a vapor of the element sodium in which the density is low enough for each atom to act independently of its neighbors. The sodium atom has 11 electrons, but 10 of them lie in closed shells, and it is only the single valence electron that takes a direct part in the interactions of the atom with light. When the sodium atom is in the ground state the outermost electron occupies an orbital designated $3S_{1/2}$. The next-highest energy levels (the next rungs on the ladder) are labeled $3P_{1/2}$ and $3P_{3/2}$, and they lie at energies 2.103 and 2.105 electron volts above the ground state. These are the smallest quantities of energy a



CHAMELEONLIKE GEMSTONE seems to adapt its color to the spectrum of the light with which it is illuminated. The stone is alexandrite, which appears red in the red- and yellow-rich light of a candle flame or an incandescent lamp but turns blue-green in sunlight or in the light from a fluorescent lamp. Alexandrite is a beryllium aluminate, BeAl_2O_4 , but the color is generated by chromium ions present in the crystal as impurities. Transitions between various energy levels of unpaired electrons in the chromium ions are responsible for the colors of several gemstones. The exact color produced depends on the

chemical environment of the ions. For example, chromium ions give rise to the red of ruby and to the green of emerald under slightly different conditions. The environment of the chromium ions in alexandrite is intermediate between that of ruby and that of emerald, and so the color is changeable. Natural alexandrites are exceedingly rare. This one is a synthetic crystal grown by Creative Crystals, Inc., of Concord, Calif. The photographs, which were made by Fritz Goro, do not show the exact colors of the crystal because the response of photographic film is somewhat different from that of the human eye.

sodium atom in the ground state can absorb. They correspond to wavelengths of 589.6 and 589.1 nanometers, in the yellow part of the spectrum.

Above the $3P$ orbitals are a multitude of other excited states, where the electron has a greater average distance from the nucleus and a higher average energy. The number of such states is infinite, but the interval between levels becomes smaller as the energy increases, so that the series converges on a finite limit. For sodium the limit comes at 5.12 electron volts, where the outermost electron is no longer merely excited but is torn loose from the atom entirely; in other words the atom is ionized.

Suppose a sodium atom is ionized,

perhaps by a quantum of ultraviolet radiation, and the free electron and the ion then recombine. Initially the electron may occupy one of the higher orbitals, but it quickly falls to a lower energy level. If the descent were made in a single step, from the ionization limit to the ground state, the atom would emit a single ultraviolet photon with an energy of 5.12 electron volts. A much likelier route would pass through several intermediate states, accompanied by the emission of a lower-energy quantum at each stage. Not all such cascades are possible; "selection rules" determine which ones are allowed. Most of the allowed pathways proceed through one of the $3P$ orbitals and thence to the ground

state. As a result quanta of yellow light with energies of 2.103 and 2.105 electron volts are among those emitted. Indeed, these two lines are by far the brightest in the spectrum of atomic sodium and a vapor of excited or ionized sodium glows bright yellow.

The characteristic yellow radiance of atomic sodium can be observed when a salt of sodium is heated in a flame hot enough to vaporize some of the atoms. In analytic chemistry this property serves as the basis of the flame test for the presence of sodium. The doublet of yellow lines is also prominent in the spectrum of a sodium-vapor lamp, where the sodium atoms are ionized by a high-voltage discharge.

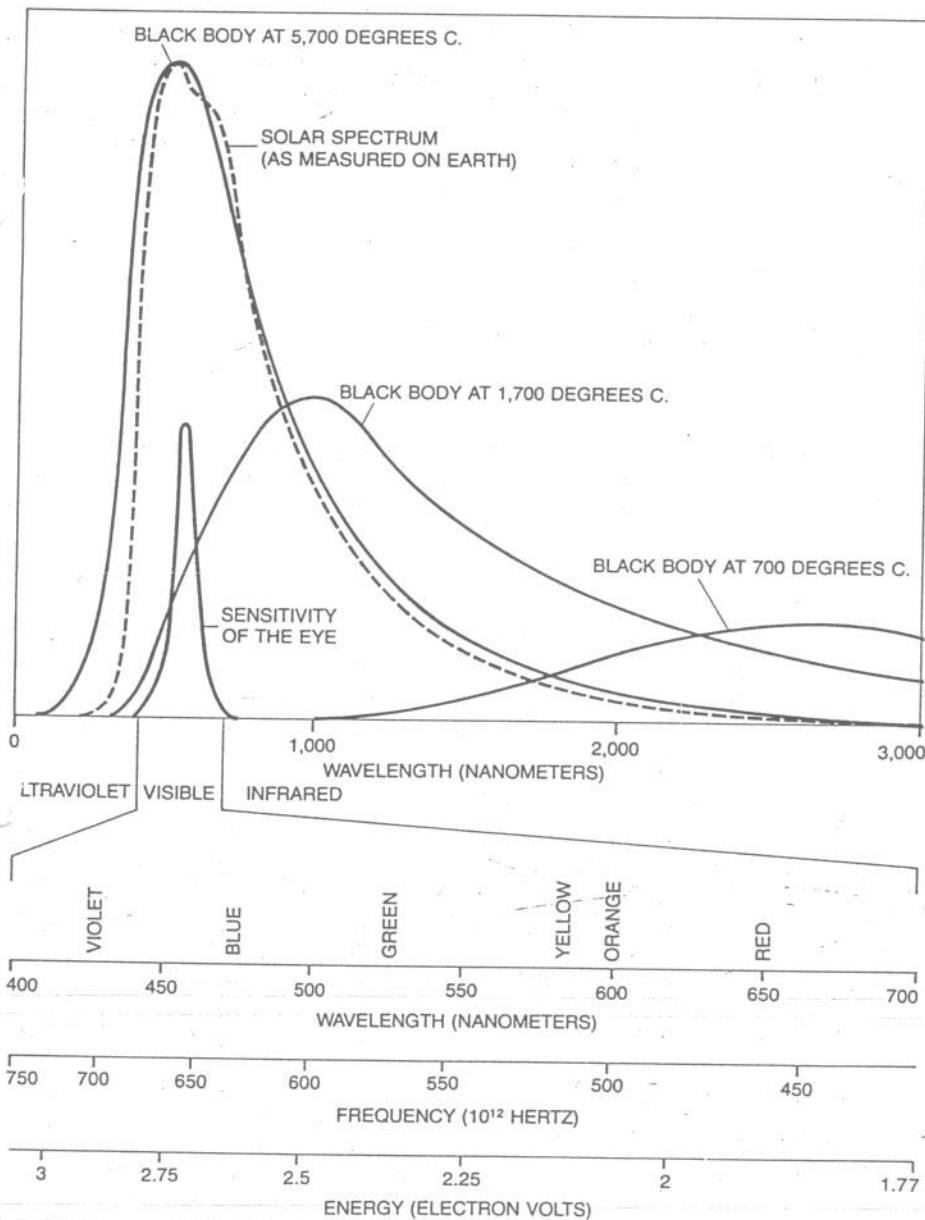
Other atoms also yield distinctive emission lines when they are excited or ionized and then allowed to return to the ground state; in each element, however, the spacing of the energy levels is different, and so the color of the emitted light also differs. In neon the strongest lines are in the red part of the spectrum, which accounts for the red glow of neon lights and signs. The mercury atom has prominent lines in the green and the violet regions of the visible spectrum, and consequently a mercury-vapor lamp gives off blue-tinted light that is deficient in red and yellow. Lasers whose working medium is a monatomic gas exploit emission lines of the same kind. Lightning and electric arcs also derive their color from electronic excitations of the atoms in gases.

Black-body Radiation

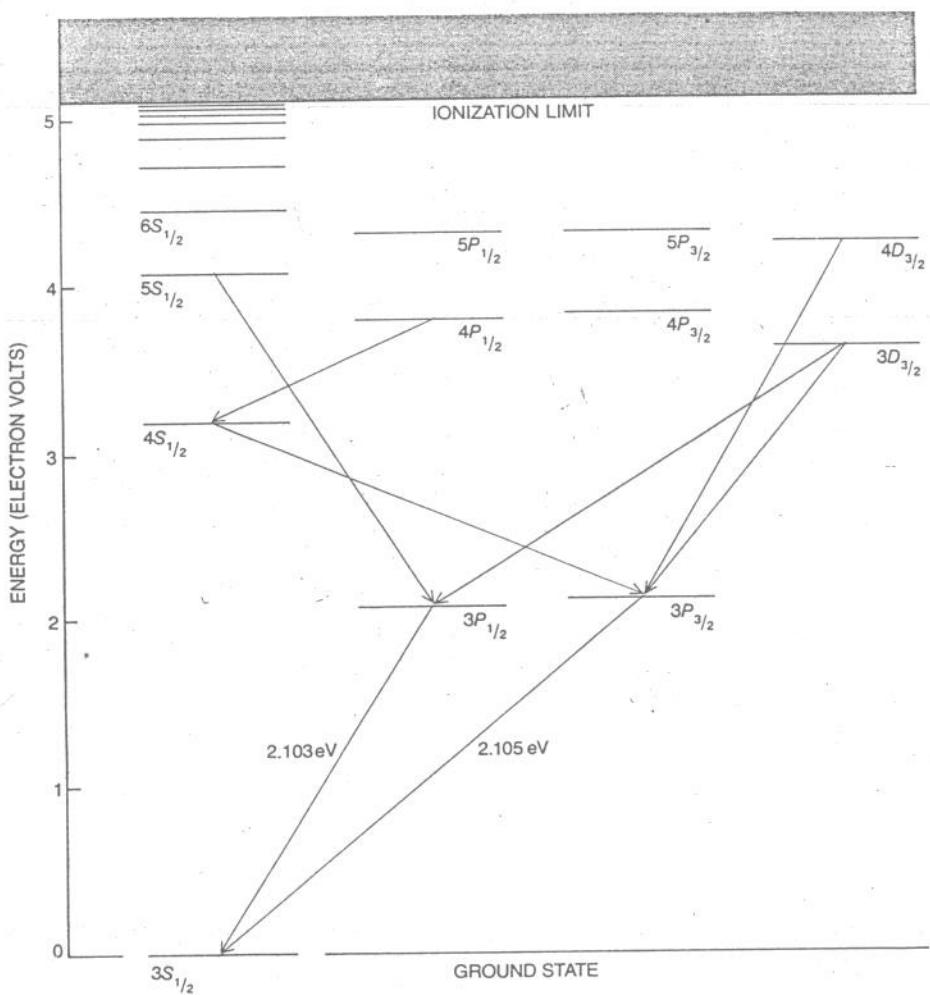
Sharply defined emission and absorption lines are typical of gases. The spectrum of light emitted by a solid or a liquid is usually quite different, in that it extends over a continuous range of wavelengths.

A universal form of radiation from condensed matter is black-body radiation, which has a continuous spectrum with a distinctive shape. Here "black body" refers simply to an idealized material that absorbs all wavelengths without favor and is also a perfect emitter of all wavelengths. (Real materials all have lower emissivity, but many approach the black-body spectrum at high temperature.) Such radiation has an important place in the history of physics, since it was through an analysis of the black-body spectrum that Max Planck deduced the quantum principle in 1900. He found he could explain the shape of the spectrum only by assuming the quantization of energy.

In ideal black-body radiation the spectrum is independent of the chemical composition of the emitter and is determined by a single parameter: temperature. At absolute zero all the atoms occupy the lowest energy level available, and no radiation is emitted. As the temperature rises some atoms are promoted



SPECTRUM OF SUNLIGHT closely matches the sensitivity of the eye; as a result light appears white if its spectrum resembles the solar one, and other colors can be described by how they depart from the solar spectrum. The sun's radiation is approximately that of a black body with a temperature of 5,700 degrees Celsius. The shape of a black-body spectrum is determined entirely by the temperature, becoming steeper and shifting to shorter wavelengths as the temperature increases. Thus as an object is heated its color changes from black (no emission) to red to yellow to white and finally to pale blue. The spectrum can be measured in units of wavelength, frequency or energy, which are merely alternative means of describing a light wave.



LADDER DIAGRAM for the sodium atom defines a spectrum of discrete wavelengths, which are the only ones the atom can emit or absorb. In order to climb to a higher rung the atom must absorb a quantum of radiation whose energy corresponds exactly to the difference in energy between the initial and final states. On falling to a lower rung the atom emits a quantum with the same energy. Most downward transitions pass through the levels designated $3P_{1/2}$ and $3P_{3/2}$ to the lowest level, or ground state, labeled $3S_{1/2}$. In these transitions quanta are emitted with energies of 2.103 and 2.105 electron volts, in the yellow part of the spectrum, and so a vapor of excited sodium atoms glows bright yellow. In the ladder diagram only the vertical dimension has meaning, but the various series of levels are separated horizontally for clarity.

to excited states, but the process is a random or statistical one, and the atoms are distributed over a broad range of energies. At any finite temperature the number of occupied states increases gradually with energy up to some maximum value; then it declines again. Thus the shape of the spectrum is somewhat like the profile of an ocean wave about to break. The steepness of the wave and the position of the crest depend on the temperature of the body.

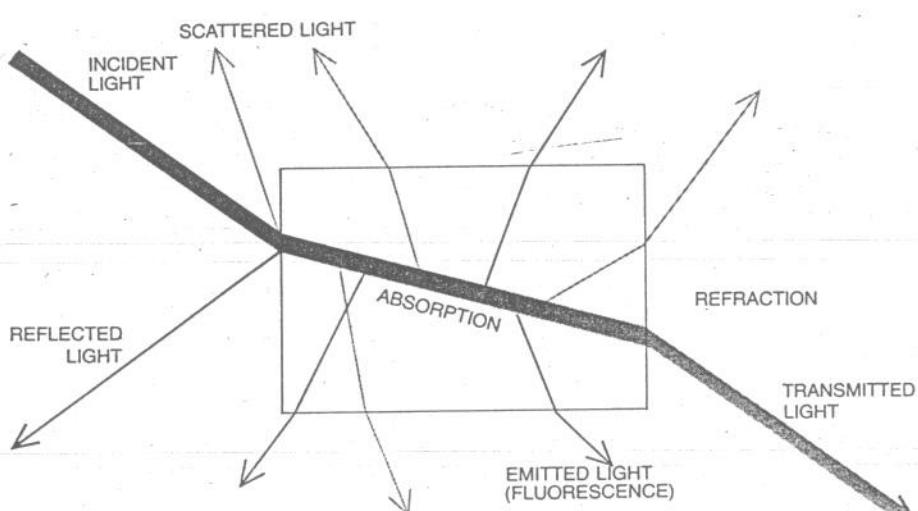
At room temperature the thermal excitations are confined to small energies, and radiation is emitted only in the infrared. When the temperature reaches about 700 degrees Celsius the maximum emissions are still in the infrared, but a little visible light begins to appear; it is perceptible as a dull red glow. As the temperature rises further the peak of the emission curve shifts to higher energies and shorter wavelengths, so that the object glows brighter and its color changes. The sequence of colors runs from red to orange to yellow to white to pale blue, in accord with the colloquial descriptions "red hot," "white hot" and so on.

In a log fire or a candle flame incandescent particles of carbon give off radiation with an effective black-body temperature of at most 1,500 degrees C., where the light ranges from red to yellow. The tungsten filament of an incandescent light bulb has a temperature of about 2,200 degrees and yields a warm yellow-white. A flash bulb, which can reach a temperature of 4,000 degrees, yields a somewhat more accurate version of white.

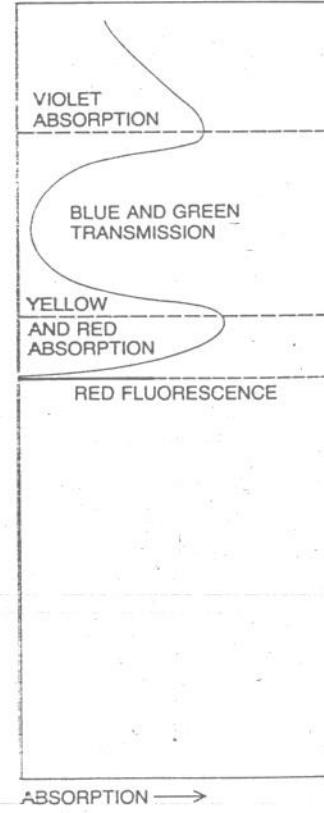
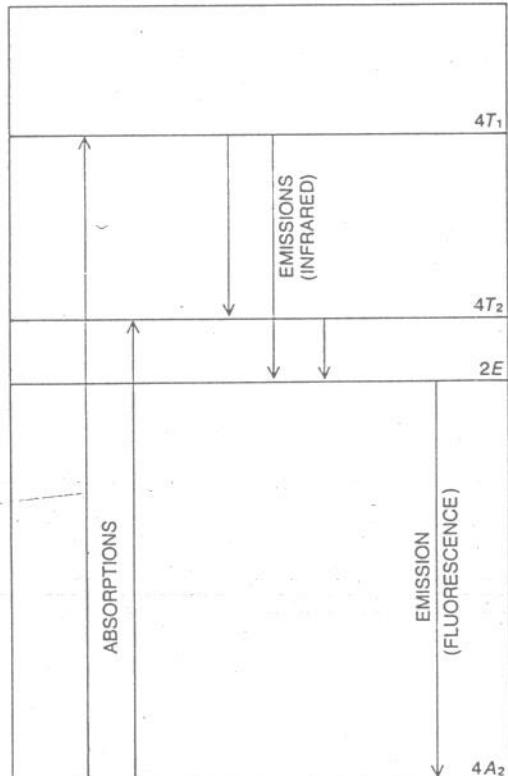
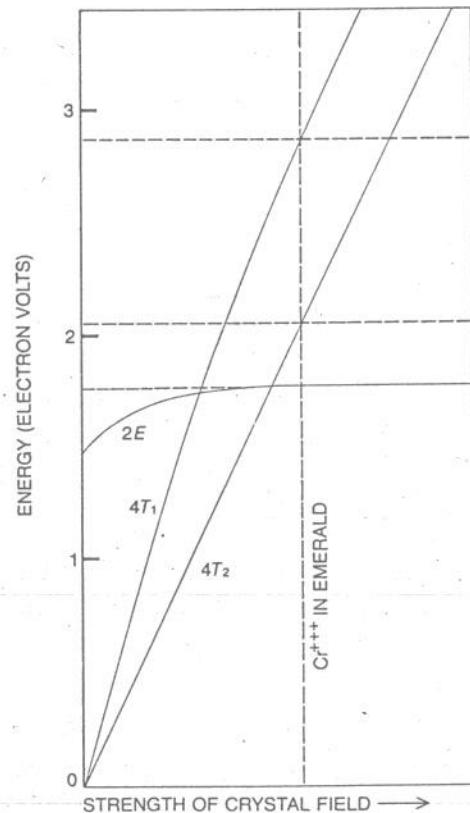
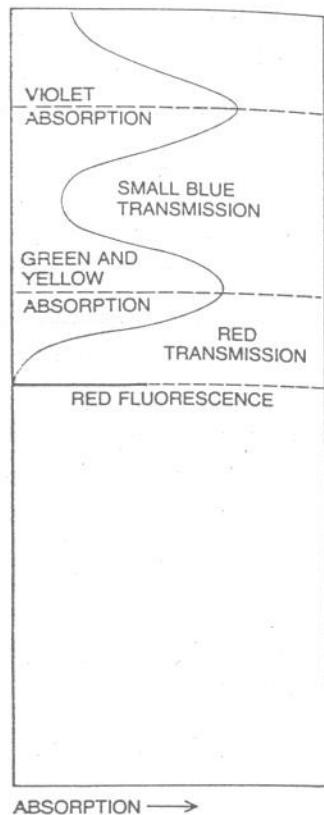
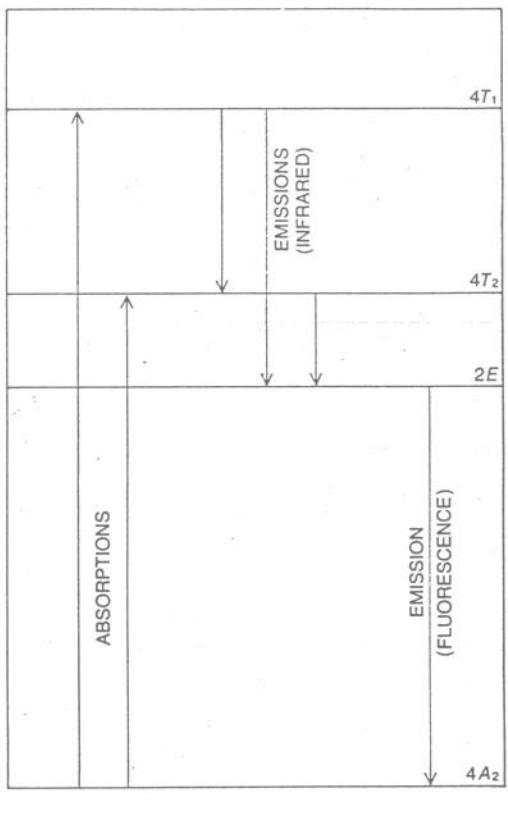
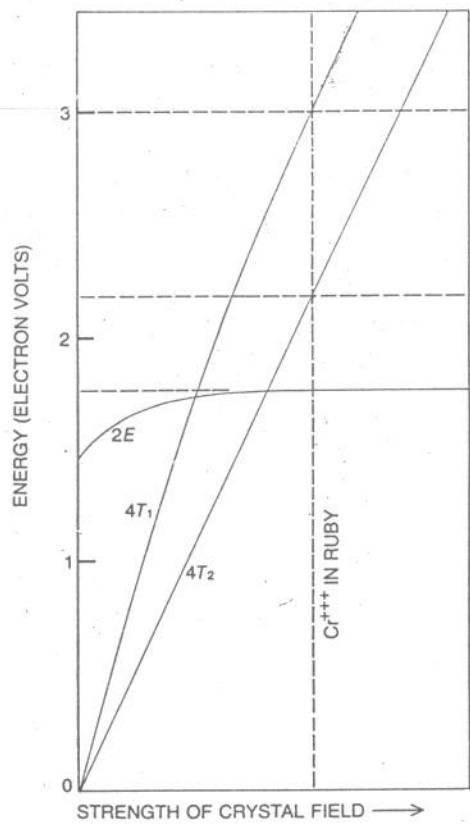
The solar spectrum has the approximate form of a black-body curve; its shape is determined by the temperature at the surface of the sun, about 5,700 degrees C. The spectrum has a broad peak centered near 2.2 electron volts, or 560 nanometers, a yellow-green wavelength. The eye is most sensitive to just this wavelength. Indeed, the concept of white seems to be conditioned largely by the spectrum of daylight, which is dominated by solar radiation. Roughly speaking, light is perceived as being white if its spectrum resembles that of sunlight; other colors can be defined according to how they depart from the solar spectrum.

Crystal-field Colors

When atoms combine to form a molecule or condense to form a liquid or a solid, new modes of excitation are introduced. Among them are mechanical vibrations and rotations that are not possible in an isolated atom. For example, the atoms of a diatomic molecule can oscillate as if they were connected by a spring, and they can rotate about their common center of mass. Such motions can occasionally influence the color of a material. In water, for example, a complex bending of the molecules absorbs a



INTERACTIONS OF LIGHT with condensed matter include reflection, refraction, scattering and absorption; some absorbed light can also be reemitted (usually at a longer wavelength) as fluorescence. The effects of each of these processes can vary with wavelength and so can give rise to color. For example, the preferential absorption of short wavelengths and reflection or transmission of long ones makes an object appear yellow, orange or red. In general condensed matter absorbs broad and essentially continuous bands of wavelengths rather than discrete lines.



ELECTRIC FIELD IN A CRYSTAL can influence color by altering the state of atoms or ions within the crystal structure. In ruby and emerald color results from the absorption of selected wavelengths by unpaired electrons in chromium ions. The transitions that cause the absorption are the same in both cases: ions are promoted from the ground state, 4A₂, to the excited levels 4T₂ and 4T₁. The energies at which these states lie are set by the magnitude of the crystal field. In ruby, the absorption bands block violet light and green and yellow

light; red light is transmitted and so is a little blue, which gives the ruby its deep red color with a slight purple cast. In emerald the crystal field is weaker, which depresses both absorption bands. As a result red transmission is eliminated and green and blue are enhanced. In both materials an excited ion returns to the ground state through an intermediate level designated 2E, whose energy is little affected by the crystal field. Transitions from the 2E level to the ground state give rise to red fluorescence that is almost identical in ruby and emerald.

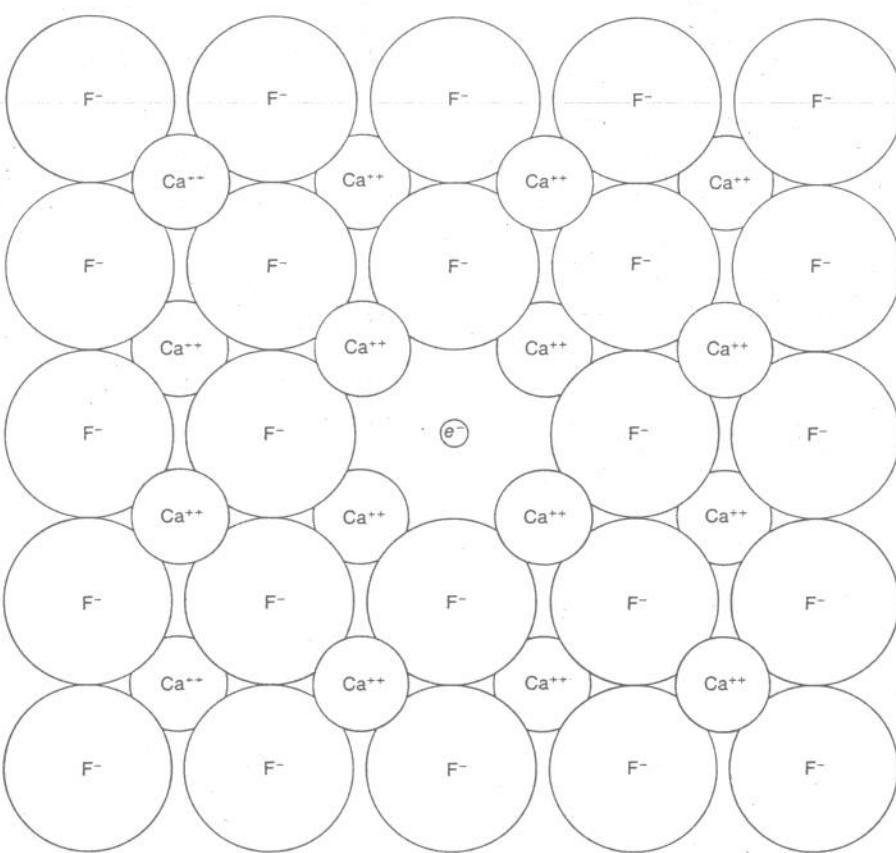
little energy at the red end of the spectrum and gives pure water and ice a pale blue cast. For the most part, however, the energy of vibrational and rotational excitations is small, and it is dissipated as infrared radiation, or heat.

Another consequence of the binding together of atoms is a change in the state of the valence electrons. In an isolated atom the valence electrons are unpaired and are the primary cause of color. In a molecule and in many solids, on the other hand, the valence electrons of one atom form pairs with the valence electrons of adjacent atoms; it is these pairs that constitute the chemical bonds that hold the atoms together. As a result of this pair formation the absorption bands of the valence electrons are displaced to ultraviolet wavelengths, and they are no longer available for the production of color. Only electrons in exceptional states remain to give rise to coloration. It is evident, however, that such exceptional states cannot be too rare; if they were, most molecules and solids would be transparent to visible light.

One set of unusual electronic states appears in the transition-metal elements, such as iron, chromium and copper and in the rare-earth elements. The atoms of metals in the transition series have inner shells that remain only partly filled. These unfilled inner shells hold unpaired electrons, which have excited states that often fall in the visible spectrum. They are responsible for a wide range of intense colors. For example, both ruby and emerald derive their color from trace amounts of chromium.

The basic material of ruby is corundum, an oxide of aluminum with the formula Al_2O_3 . Pure corundum is colorless, but in ruby a brilliant color results from the substitution of chromium ions (Cr^{+++}) for a few percent of the aluminum ions. Each chromium ion has three unpaired electrons, whose lowest-possible energy is a ground state designated $4A_2$; there is also a complicated spectrum of excited states. All the excited states are broadened to form bands, and they are also modified in another way by the presence of the crystal matrix. Although the identity of the states is defined by the electronic configuration of the chromium ion, the absolute position of each level in the energy spectrum is determined by the electric field in which the ion is immersed. The symmetry and strength of the field are determined in turn by the nature of the ions surrounding the chromium and by their arrangement.

In ruby each chromium ion lies at the center of a distorted octahedron formed by six oxygen ions. The interatomic bonds in ruby are about .19 nanometer long, and they have about 63 percent ionic character, which means that the electron pairs that make up the bonds spend more of the time near the oxygen ions than they do near the aluminum or



COLOR CENTER can form in a crystal when an electron takes the place of a dislodged ion. In fluorite, or calcium fluoride, the electron fills the vacancy created when a fluorine ion is removed. The electron has a spectrum of excited states that extends into the range of visible wavelengths. The color centers in fluorite, which are called F centers, give it a purple hue. The original crystal structure can be restored by heating, whereupon the color fades. There are also electron and "hole" color centers, where a single electron rather than an entire ion is displaced.

chromium ions. This distribution of the electronic charge gives rise to a comparatively strong electric field, which is called the crystal field or the ligand field. When a chromium ion is immersed in this field, three excited states of its unpaired electrons have energies in the visible range.

The three excited states are designated $2E$, $4T_2$ and $4T_1$. Selection rules forbid a direct transition from the ground state to the $2E$ level, but both of the $4T$ levels can be entered from the ground state. The energies associated with these transitions correspond to wavelengths in the violet and in the yellow-green regions of the spectrum. Because the levels are not sharp lines but broad bands a range of wavelengths can be absorbed. Hence when white light passes through a ruby it emerges depleted of its violet and yellow-green components. Essentially all red is transmitted, along with some blue, giving the ruby its deep red color with a slight purple cast.

Because of the selection rules electrons can return from the excited $4T$ levels to the $4A_2$ ground state only through the intermediate $2E$ level. The initial transitions from $4T$ to $2E$ release small amounts of energy corresponding to in-

frared wavelengths, but the drop from $2E$ to the ground state gives rise to strong emission of red light. It should be noted that this red light, unlike the transmitted bands, is not present in the beam originally incident on the crystal; it is generated within by the process of fluorescence. Indeed, the red fluorescence of ruby can be observed most clearly when the crystal is illuminated with green or violet light or with ultraviolet radiation. On the other hand, the fluorescence can be quenched by iron impurities, which are often present in natural rubies. The light of a ruby laser derives from red fluorescence in synthetic rubies, which are free of iron.

Ruby and Emerald

The subtlety of crystal-field colors can be made apparent through a comparison of ruby and emerald. The color-generating impurity in emerald is again the Cr^{+++} ion, and it again replaces aluminum in small amounts. The similarity of the two substances extends further: in emerald too the chromium ions are surrounded by six oxygen ions in an octahedral configuration, and the bond length is again about .19 nanometer. In emerald,

however, the fundamental crystal lattice is that of a beryllium aluminum silicate, $\text{Be}_3\text{Al}_2\text{Si}_6\text{O}_{18}$. And the most significant difference is in the nature of the chemical bonds, which are less ionic by a few percent, so that the magnitude of the electric field surrounding a chromium ion is somewhat reduced. As a result the two $4T$ levels lie at slightly lower energies; the position of the $2E$ band is essentially unaltered. The major effect of these changes is to shift the absorption band that in ruby blocks green and yellow light downward into the yellow and red part of the spectrum. The emerald therefore absorbs most of the red light, but the transmission of blue and green is greatly enhanced.

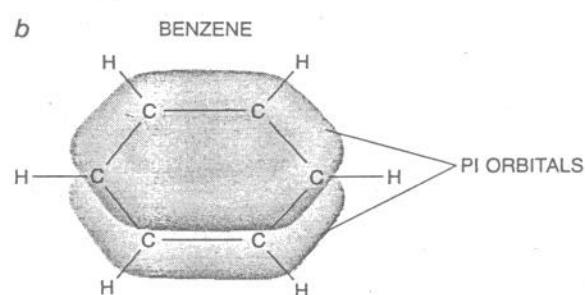
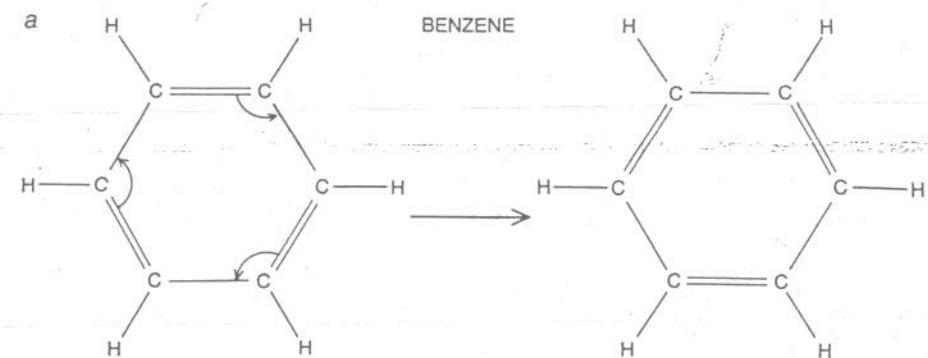
Curiously fluorescence in emerald is almost identical with that in ruby. The reason is that the energy of the $2E$ level is scarcely altered by the reduction in the crystal field. The similarity of the fluorescent emissions shows it is merely a coincidence that ruby has a red color and also red fluorescence.

Intermediate in spectral composition between ruby and emerald is the rare (and therefore costly) gemstone alexandrite. Again the color arises from chromium ions that replace aluminum, but in this case the underlying crystal structure is that of a beryllium aluminate, BeAl_2O_4 . The crystal field that sets the energy scale of the chromium ions is stronger than that in emerald but weaker than that in ruby, with the result that the red and green transmission bands are quite evenly balanced. The near equality of the two bands has an extraordinary consequence: in blue-rich sunlight the gemstone appears blue-green, but in the redder light of a candle flame or an incandescent lamp it appears red.

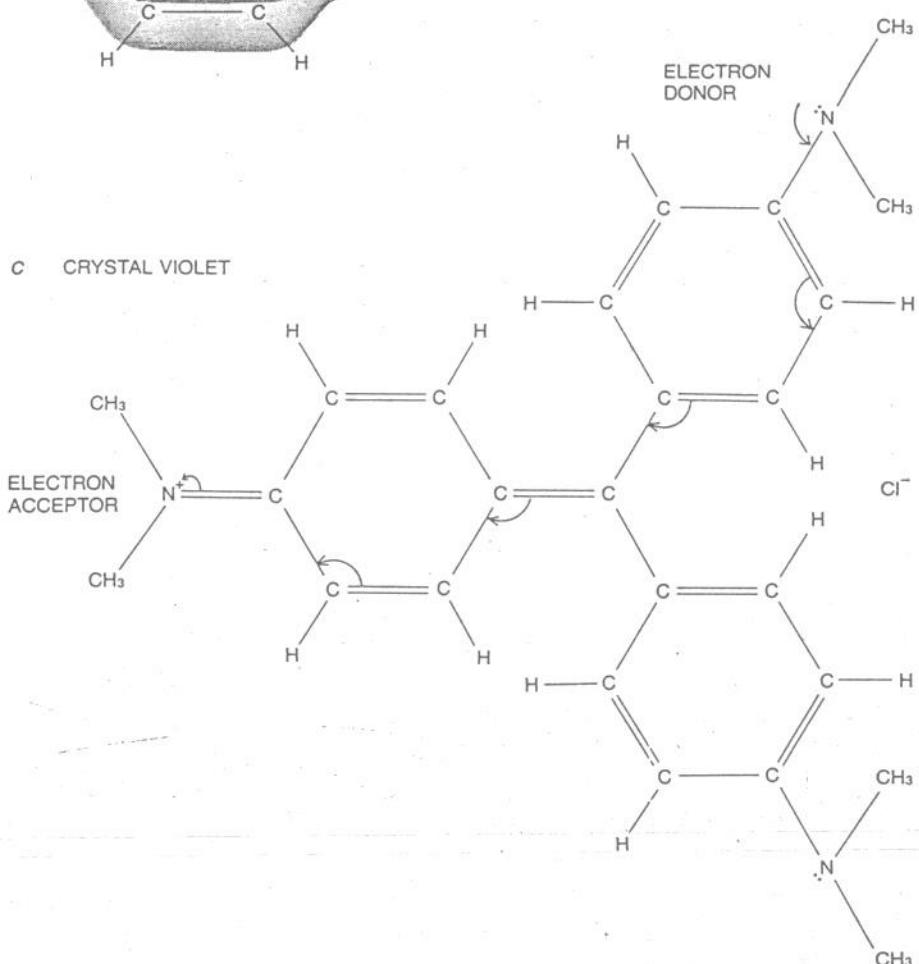
Crystal-field colors can arise whenever there are ions bearing unpaired electrons in a solid. Aquamarine, jade and fine quartz get their colors from a mechanism similar to that in ruby and emerald, but the transition-metal impurity is iron instead of chromium. Many compounds and crystals in which the transition metals appear as major constituents rather than impurities are also strongly colored. Among natural minerals in this category are the blue or green azurite, turquoise and malachite, in which the color is produced by copper, and the red garnets, which owe their color to iron. Most pigments in paints are transition-metal compounds.

Color Centers

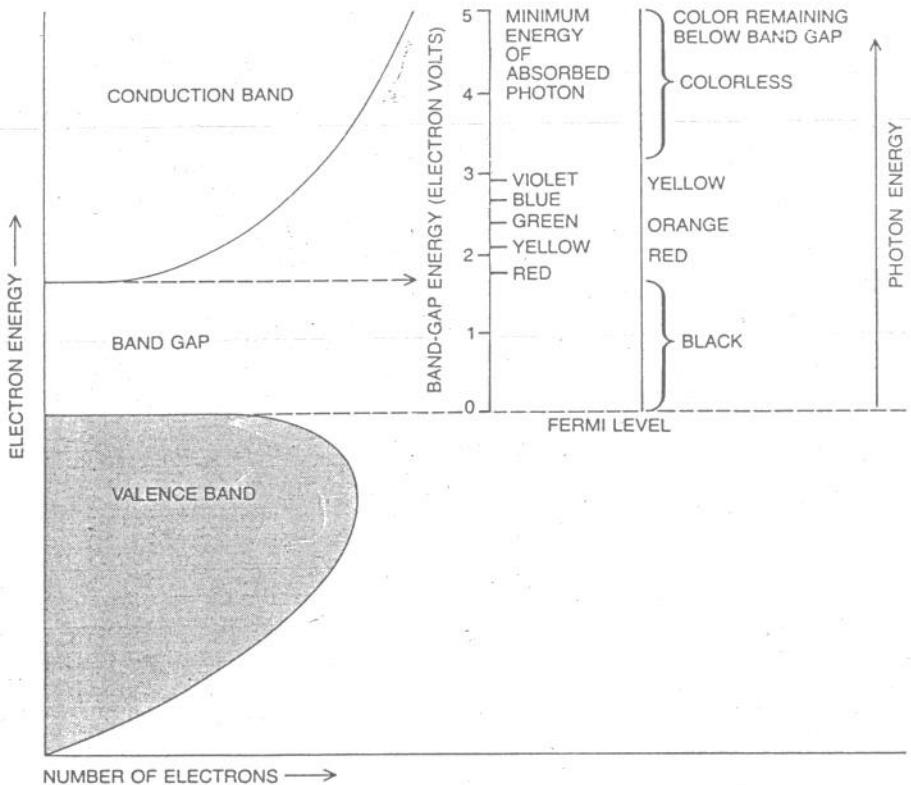
The physical mechanism responsible for crystal-field colors is not confined to electrons in transition-metal ions; indeed, the electrons need not be an intrinsic component of any atom. An excess electron unattached to any single atom will suffice if the electron can be trapped at some structural defect, such as a miss-



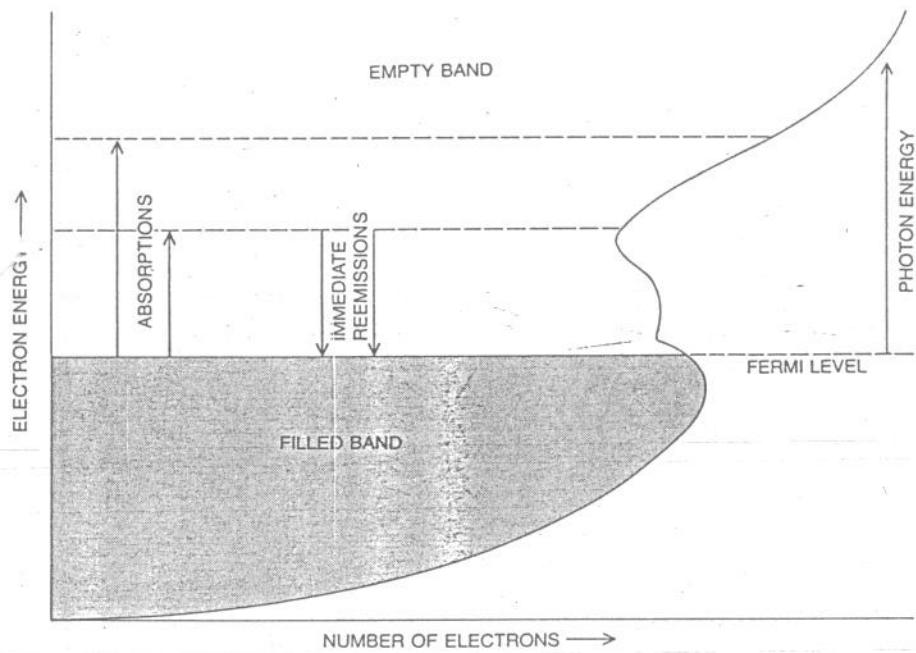
c CRYSTAL VIOLET



MOLECULAR ORBITALS describe paired electrons distributed over several atoms or over many atoms; they often absorb strongly at visible wavelengths. The pair of electrons that forms a chemical bond ordinarily absorbs only in the ultraviolet, but when alternative configurations of the bonds spread the pair over a number of atoms, the excitation energy of the pair is reduced. The commonest molecular orbitals that cause color are those associated with systems of conjugated bonds (alternating single and double bonds) in organic compounds such as benzene (a). By shifting three pairs of electrons in benzene the sequence of bonds is reversed. A better way of representing the structure of the molecule (b) shows single bonds between all the carbon atoms, with the extra three pairs of electrons distributed throughout the molecule in "pi" orbitals. It is the pi orbitals that are active in producing color. In benzene the excited states of the pi orbitals still lie in the ultraviolet, and so benzene is colorless, but in molecules with larger conjugated systems, such as the dye crystal violet (c), the bands are at visible wavelengths. The color is enhanced by chemical groups called auxochromes that donate and accept electrons.



ELECTRONIC STRUCTURE OF A METAL is distinguished by an essentially continuous band of allowed energy levels. The band is filled from the ground state up to an energy called the Fermi level; all higher energy states are empty and can therefore accept excited electrons. A consequence of this electronic configuration is that all wavelengths of radiation can be absorbed, from the infrared through the visible to the ultraviolet and beyond. A material that absorbs light of all colors might be expected to appear black; metals are not black because an excited electron can immediately return to its original state by reemitting a quantum with the same wavelength as the absorbed one. The metallic surface is therefore highly reflective.



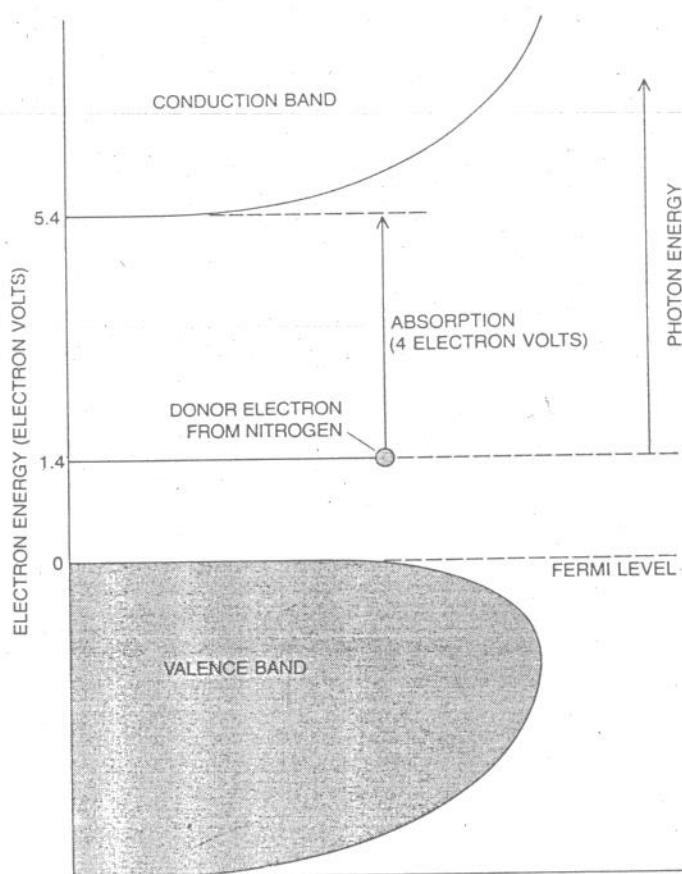
BAND STRUCTURE OF A SEMICONDUCTOR is similar to that of a metal, except that a gap of forbidden energies separates the filled valence band from the empty conduction band. There is therefore a minimum energy that a quantum of radiation must have in order to be absorbed, namely the energy needed to promote an electron from the top of the valence band to the bottom of the conduction band. The color of a pure semiconductor is determined by the magnitude of the band gap. If it lies in the infrared, all visible wavelengths are absorbed and the material is black. A gap energy in the visible region allows some colors to be transmitted, so that the semiconductor takes on a color ranging from red to yellow. When the gap energy is in the ultraviolet, all visible wavelengths are transmitted and the material is colorless.

ing ion or an impurity. A "hole," the absence of one electron from a pair, can have the same effect. Anomalies of this kind are called color centers or *F* centers (from the German *Farbe*, color).

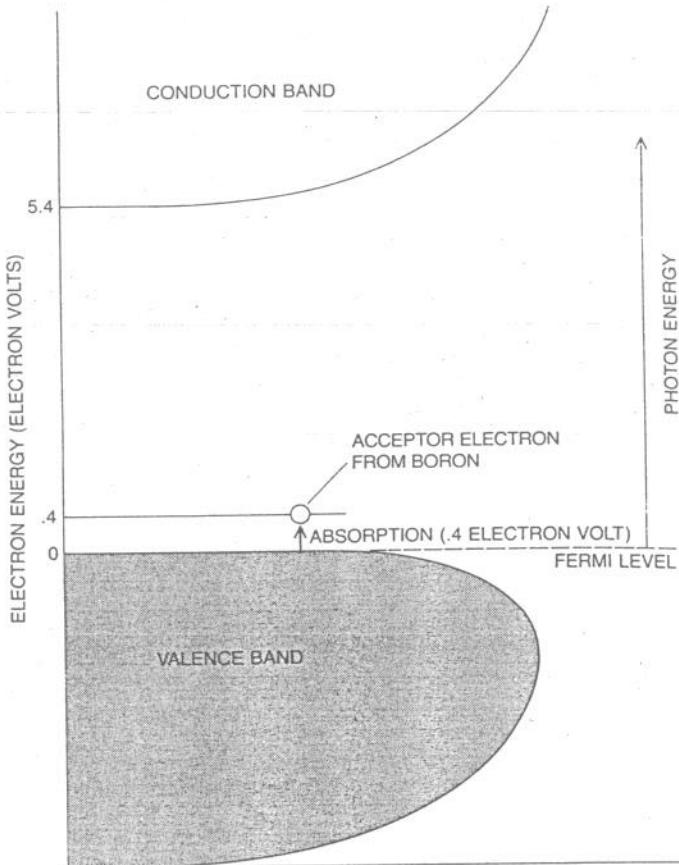
Although many color centers are known, the mechanism of color-production is understood in only a few. One of these is the purple *F* center of fluorite, a mineral that exhibits a great variety of color-center activities. Fluorite is calcium fluoride, or CaF_2 , in which each calcium ion is normally surrounded by eight fluorine ions. An *F* center forms when a fluorine ion is missing from its usual position. The loss of the fluorine can come about in several ways: by growing the crystal in the presence of excess calcium, by exposing the crystal to high-energy radiation (which can displace an ion from its usual position) or by applying a strong electric field (which removes fluorine by the process of electrolysis). In order to preserve the electrical neutrality of the crystal some other negatively charged entity must take up the position in the lattice left vacant by the absent fluorine ion. When the charge is supplied by an electron, an *F* center is created. The electron is bound in place not by a central nucleus, as in an atom or an ion, but by the crystal field of all the surrounding ions. Within the field it can occupy a ground state and various excited states similar to those in the transition metals. The movement of electrons between these states gives rise both to color and to fluorescence.

The color of smoky quartz is attributed to a hole color center. The basic lattice of quartz is silicon dioxide (SiO_2), but a prerequisite to the formation of the color center is the presence of aluminum impurities replacing a few silicon ions. Because aluminum has a valence of +3 and silicon a valence of +4, an alkali metal ion or a hydrogen ion must be nearby to maintain electrical neutrality. Quartz is almost always contaminated with traces of aluminum, but that alone will not give rise to color because there are no unpaired electrons. The color center is created when the quartz is exposed for a few minutes to intense X rays or gamma rays, or when it is exposed to low levels of such radiation over geologic periods. The radiation expels one electron from a pair of electrons in an oxygen atom adjacent to an aluminum impurity, thereby leaving an unpaired electron in the orbital. The absent electron is called a hole, and the remaining unpaired electron has a set of excited states much like that of an excess electron.

Amethyst is another form of quartz that derives its color from a hole color center, but the impurity ion is iron rather than aluminum. Certain old bottles were made of a glass containing iron and manganese; after many years of exposure to intense sunlight the glass turns purple through the development of col-



DOPING OF A SEMICONDUCTOR with impurities creates allowed energy levels within the band gap. In diamond, which is a semiconductor with a band gap of 5.4 electron volts, doping with nitrogen introduces a level of filled states 1.4 electron volts above the Fermi level. One of these electrons can be excited to the conduction band by absorbing radiation with an energy of four electron volts. Because



the level is broadened somewhat, there is some absorption of violet and blue light, and the diamond is colored yellow. Adding boron to a diamond creates holes, or empty states, in a level centered .4 electron volt above the Fermi level. Light of the lowest visible energies can promote an electron from the valence band into this broadened level, so that red and yellow light is absorbed and the diamond becomes blue.

or centers and is called desert-amethyst glass. The same effect is achieved by a 10-minute exposure to the intense gamma rays emitted by cobalt 60.

Most color centers are stable if the material is not heated excessively. In fluorite raising the temperature makes the displaced fluorine ion mobile, so that it eventually resumes its original position; the color center and the color are then abolished. Amethyst when heated changes color, becoming either citrine quartz or a rare greened quartz. Both these colors represent the influence of iron without the amethyst color center. In some materials even the energy of sunlight can cause a color center to fade. For example, some colorless topaz can be irradiated to induce a deep orange-brown tint; the color fades, however, after only a few days in sunlight. Natural topaz of the same color is quite stable.

Molecular Orbitals

It was pointed out above that in molecules and solids the valence electrons are paired in chemical bonds and as a result their excited states are shifted into the ultraviolet. Actually this is true only

when the paired electrons remain confined to a particular bond between two atoms. In many cases the electrons can move over longer distances; they can range throughout a molecule or even throughout a macroscopic solid. They are then bound less tightly, and the energy needed to create an excited state is reduced. The electrons are said to occupy molecular orbitals (in contrast to atomic orbitals), and they are responsible for a varied class of colors in nature.

One mechanism by which molecular orbitals can contribute color is the transfer of electric charge from one ion to another; blue sapphire provides an example of the process. Like ruby, sapphire is based on corundum, but it has two significant impurities, iron and titanium, both appearing in positions normally filled by aluminum. In the lowest energy state the iron has a formal charge of +2 and the titanium has a formal charge of +4. An excited state is formed when an electron is transferred from the iron to the titanium, so that both ions have a formal charge of +3. An energy of about two electron volts is needed to drive the charge transfer. Such transitions create a broad absorption band

that extends from the yellow through the red and leaves the sapphire with a deep blue color.

In a number of materials iron is present in both its common valences, Fe^{++} and Fe^{+++} . Charge transfers between these forms give rise to colors ranging from deep blue to black, as in the black iron ore magnetite.

Molecular-orbital theory also applies to the colors observed in many organic substances in which carbon atoms (and sometimes nitrogen atoms) are joined by a system of alternating single and double bonds, which are called conjugated bonds. The best-known example of such a conjugated system is the six-carbon benzene ring, but there are many others. Because each bond represents a pair of shared electrons, moving a pair of electrons from each double bond to the adjacent single bond reverses the entire sequence of bonds. The two structures defined in this way are equivalent, and there is no basis for choosing between them. Actually the best representation of the structure shows all the atoms connected only by single bonds; the remaining pairs of bonding electrons are distributed over the entire structure

in molecular orbitals, which in this instance are called pi orbitals.

The extended nature of the pi orbitals in a system of conjugated bonds tends to diminish the excitation energy of the electron pairs. In benzene the energy of the lowest excited state is still in the ultraviolet, and so benzene is colorless, but in larger molecules, and notably in those built up of multiple rings, absorption can extend into the visible region. Such colored organic molecules are called chromophores (meaning color-bearers). Chemical side groups that donate or accept electrons can be attached to the conjugated system to enhance the color; they are called auxochromes (color-increasers).

A number of biological pigments owe their color to extended systems of pi orbitals. Among them are the green chlorophyll of plants and the red hemoglobin of blood. Organic dyes employ the same mechanism. The *Color Index* of the Society of Dyers and Colourists lists 8,000 such substances.

Some chemical species with extended molecular orbitals are also capable of fluorescence. As in ruby and emerald, the fluorescent light is emitted when an excited state decays through an intermediate level and the energy of at least one of the intermediate transitions corresponds to a visible wavelength. The fabric brighteners added to some detergents achieve their effect by absorbing ultraviolet radiation in daylight and reemitting a part of the energy as blue light. Lasers that have a dye as their active medium operate on the same principle.

Fluorescence is also observed in some systems of molecular orbitals where the energy to create the initial excited state comes from a source other than radiation. The bioluminescence of fireflies and of some deep-sea fishes is driven by a sequence of chemical reactions that culminates in the formation of an excited state of a molecule with extended pi orbitals. Manmade "cold light" devices imitate the biological process.

Metals and Semiconductors

The spatial extent of electron orbitals reaches its maximum possible value in metals and semiconductors. Here the electrons are released entirely from attachments to particular atoms or ions and can move freely throughout a macroscopic volume. Their range is limited only by the dimensions of the material. The enormous numbers of mobile electrons (on the order of 10^{23} per cubic centimeter) give metals and semiconductors optical and electrical properties unlike those of any other material.

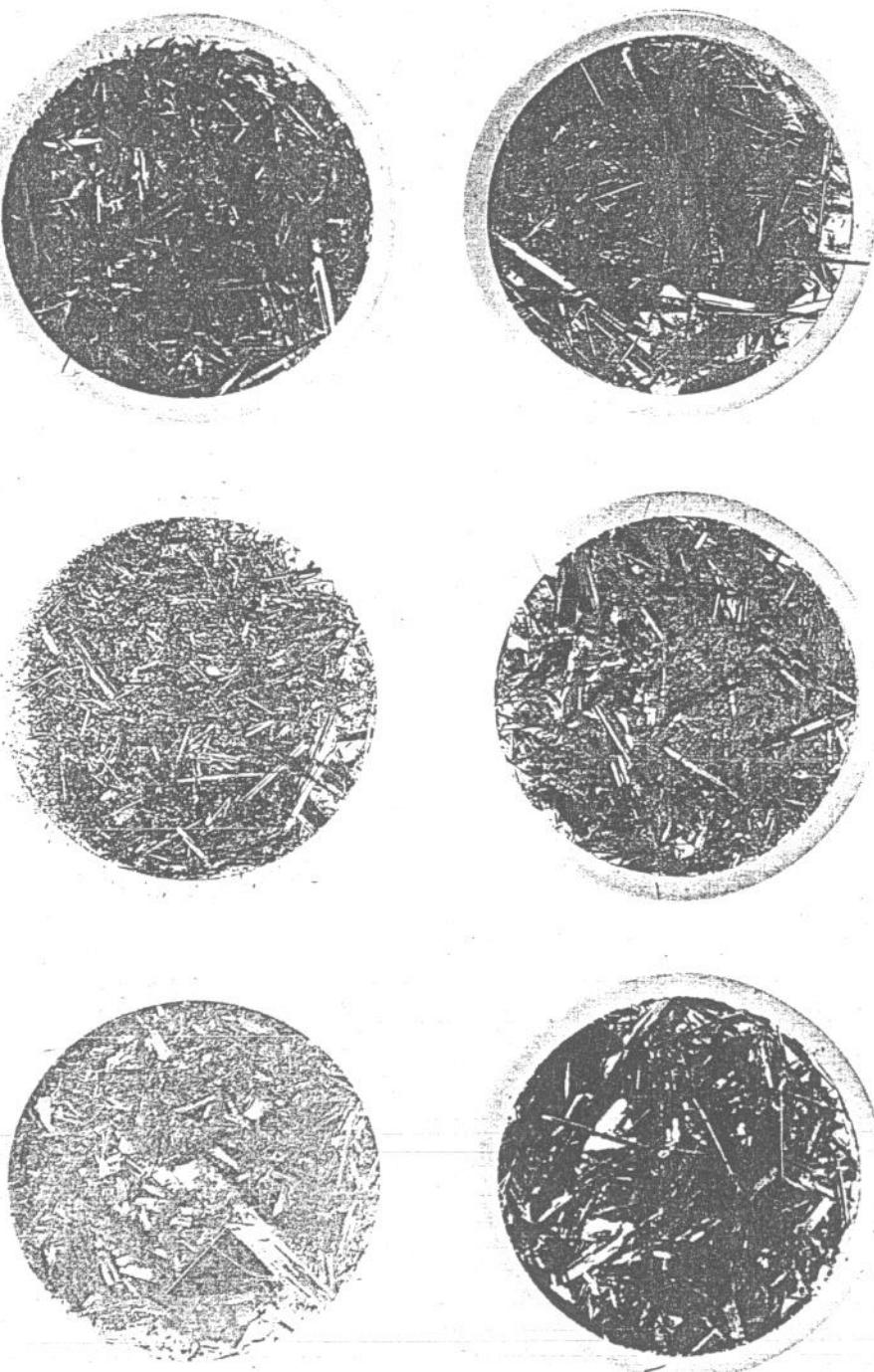
In a metal all the valence electrons are essentially equivalent since they can freely exchange places. One might suppose, therefore, that they would all have the same energy, but a rule of quantum

mechanics forbids that. In a solid as in an isolated atom only a limited number of electrons can occupy the same rung of the energy ladder. Accordingly there must be many energy levels in the metal, and the levels are necessarily spaced very closely. In effect they form a continuum from the ground state up.

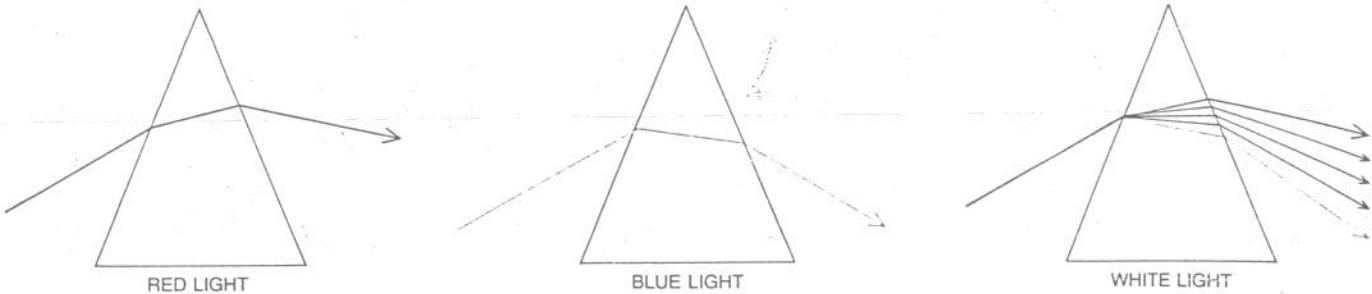
At zero temperature this continuous band of states is filled from the lowest

level up to an energy designated the Fermi level; all states above the Fermi level are vacant. Any input of energy, no matter how small, propels an electron into one of the empty states of higher energy.

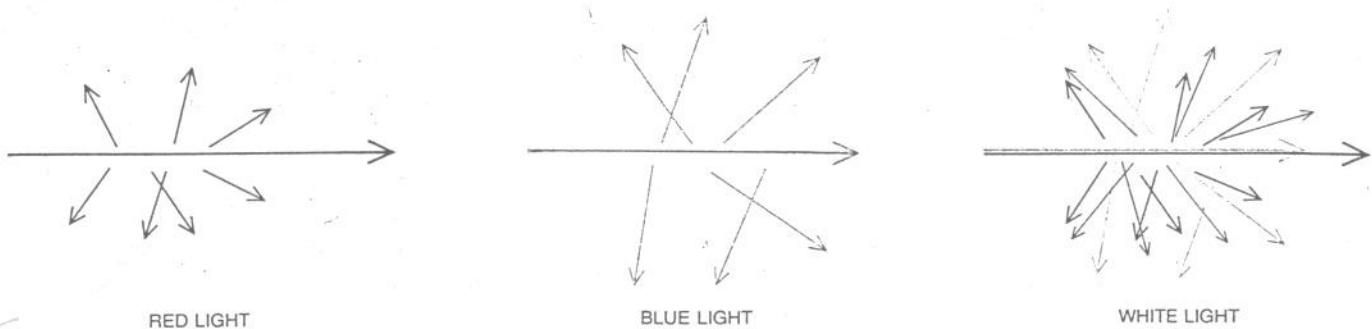
Since a metal effectively has a continuum of excited states, it can absorb radiation of any wavelength. A surface with this property of absorbing all colors might be expected to appear black, but



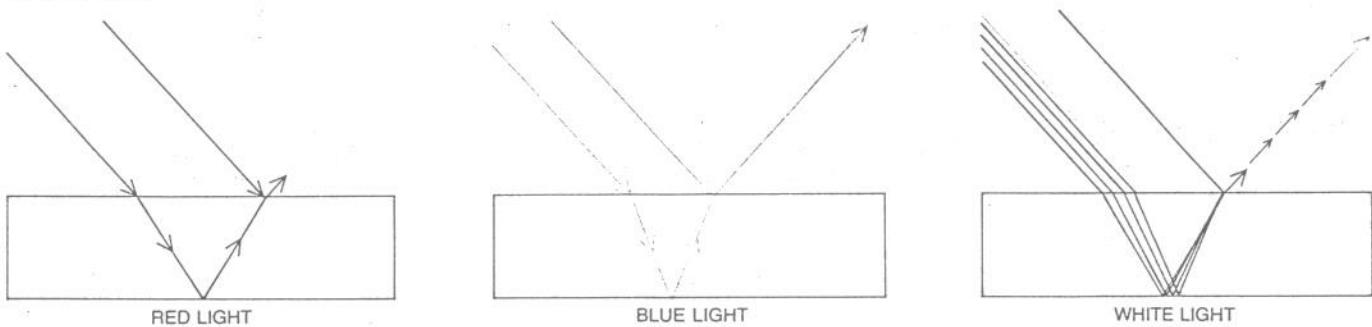
BAND-GAP COLORS are observed in six semiconducting compounds. Pure cadmium sulfide (bottom left) has a band-gap energy of 2.6 electron volts, with the result that only violet light can be absorbed; the material has the color complementary to violet, namely yellow. In pure cadmium selenide (bottom right) the band-gap energy is 1.6 electron volts, so that all visible wavelengths are absorbed and the crystals are black. The other four materials are mixtures of various proportions of the two compounds; they show a gradient in absorption, with colors from orange through red. The compounds are employed as pigments under the name cadmium red.



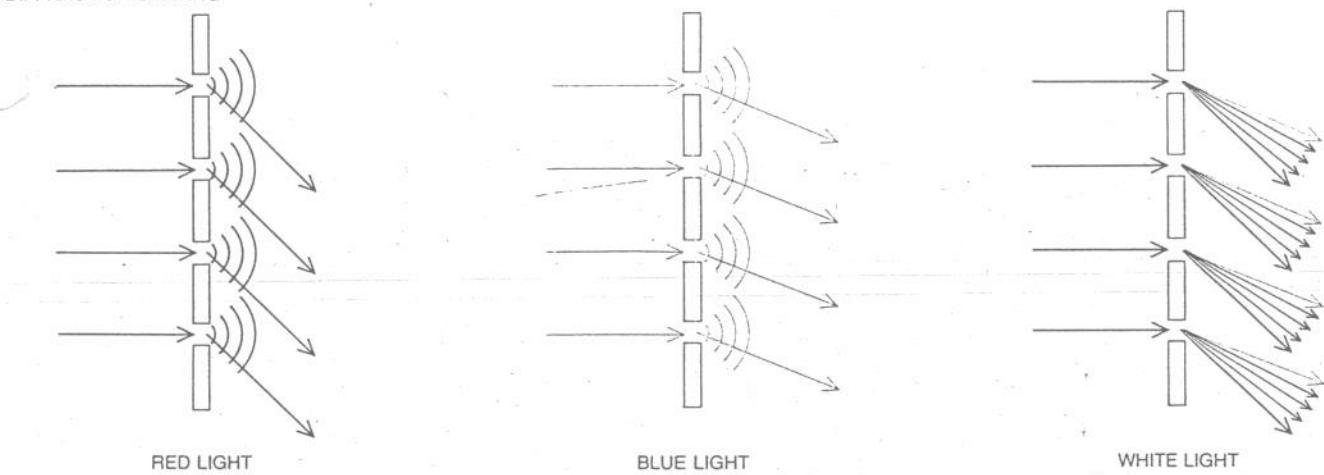
SCATTERING



INTERFERENCE



DIFFRACTION GRATING



PHYSICAL OPTICS provides the most convenient interpretation of colors generated by several mechanisms that involve a change in the direction of light. The dispersion of white light into its component colors by a transparent prism comes about because short wavelengths are refracted through a larger angle than long wavelengths. In a similar way the scattering of light by small particles is more effective at short wavelengths, so that more blue light is scattered than red. Interference is observed when a light wave is split into two parts that are

then brought together again. If the waves are in phase when they recombine, the intensity is enhanced; if they are out of phase, the waves cancel each other. Since the phase difference between the two beams can depend on wavelength, interference can enhance some colors and suppress others. In a diffraction grating light is scattered by many uniformly spaced centers and the resulting multiple wave fronts interfere with one another. Each wavelength is reinforced in one set of directions and canceled in all others, so that white light is dispersed.

metals are not an absorptive black. The reason again has to do with the agility of the metallic electrons: when an electron in a metal absorbs a photon and jumps to an excited state, it can immediately reemit a photon of the same energy and return to its original level. Because of the rapid and efficient reradiation the surface appears reflective rather than absorbent; it has the luster characteristic of metals. If the surface is smooth enough, the reflection can be specular, as in a mirror.

The variations in the color of metallic surfaces, which, for example, distinguish gold from silver, result from differences in the number of states available at particular energies above the Fermi level. Because the density of states is not uniform some wavelengths are absorbed and reemitted more efficiently than others. Besides absorption and reemission, transmission is also possible in some metals, although it is always weak and can be observed only in thin layers. The transmission varies with

wavelength, and so it gives rise to color. Beaten gold leaf, in which the gold structure is highly distorted, transmits only green light. Colloidal gold in ruby glass is not under strain, and the light transmitted is purple-red.

Band-gap Colors

Wide, continuous bands of electronic states are a feature of another class of materials: the semiconductors. The definitive characteristic of these materials is that the average number of bonding electrons per atom is exactly four. Included are crystalline forms of some elements in Group IV of the periodic table, such as silicon, germanium and the diamond phase of carbon; there are also many compound semiconductors, such as gallium arsenide.

What distinguishes a semiconductor from a metal is a splitting of the band of energy levels into two parts. All the low-energy levels form a valence band, which in the ground state is completely filled. All excited states lie in a separate conduction band, which in the ground state is entirely empty. Separating the two bands is a gap of forbidden energies. The division of the energy band has profound effects on the optical properties of a semiconductor. No longer can the electrons absorb radiation of arbitrarily low energy; the minimum energy is the energy needed to lift an electron from the top of the valence band to the bottom of the conduction band.

The color of a pure semiconductor depends only on the magnitude of the energy gap. If the gap is smaller than the lowest energy of visible light, all visible wavelengths are absorbed. Small-gap semiconductors in which reemission is efficient and rapid have a metal-like luster, as silicon does, but other small-gap

semiconductors are black. At the opposite extreme the band gap can be greater than the highest energy of visible light; in that case no visible wavelengths can be absorbed and the material is colorless. Diamond, with a band gap of 5.4 electron volts, is one such large-gap semiconductor and is transparent both to visible light and to a limited range of ultraviolet radiation.

Where the band-gap energy falls in the visible range, the semiconductor has a definite color. The mercury ore cinnabar (which has the formula HgS and is also known as the pigment vermillion) has a band-gap energy of 2.1 electron volts. All photons with energies higher than this level are absorbed, and only the longest visible wavelengths are transmitted; as a result cinnabar appears red. The pigment cadmium yellow, CdS , has a band gap of 2.6 electron volts and absorbs only blue and violet light; after these wavelengths are subtracted from white light the color remaining is yellow. The sequence of band-gap colors, from the smallest gap to the largest, is black, red, orange, yellow, colorless.

Although large-gap semiconductors are colorless when they are pure, they can take on color when they are "doped" with traces of an impurity. The impurity can be either a donor of electrons or an acceptor of them; in either case it introduces a set of energy levels in the gap between the valence band and the conduction band. Transitions to or from the new levels require a smaller quantity of energy than transitions across the entire gap.

In diamond the replacement of a few carbon atoms with nitrogen creates a donor band about 1.4 electron volts above the valence band. Nitrogen has one electron more than carbon, and it is these extra electrons that form the donor band. The nominal energy needed to promote one of them to the conduction band is four electron volts, which is still in the ultraviolet; the donor level is sufficiently wide, however, for some violet light to be absorbed. At a concentration of one nitrogen atom per 100,000 carbon atoms diamond is yellow; with more nitrogen it becomes green.

Boron has one electron fewer than carbon and gives rise to acceptor, or hole, levels in the band gap of diamond. An electron can be excited from the valence band to occupy one of these holes. The impurity level is centered .4 electron volt above the valence band but extends far enough to absorb some of the longer visible wavelengths. The boron-doped diamond therefore appears blue.

It is well known that doping also alters the electrical properties of semiconductors, making possible all the devices of solid-state electronics. Among these devices are light-emitting diodes and semiconductor lasers, in which an elec-

tric current populates excited states and the electrons emit radiation in returning to the ground state. Doped semiconductors can also function as phosphors, which are materials that give off light with high efficiency when they are stimulated electrically or by some other means. Phosphors are the luminous sources in a fluorescent lamp and in the picture tube of a television set. In a color picture tube three phosphors are distributed over the surface, with emissions at red, green and blue wavelengths. Certain compounds of the transition metals also act as phosphors.

Geometrical Optics

In the phenomena I have described so far color results either from the direct emission of colored light or from the selective absorption of some wavelengths and transmission of others. In a last group of color-producing phenomena interactions of light with matter change the direction of the light. The change in direction is the primary cause of color in refraction and diffraction, where the magnitude of the deflection can vary with wavelength. In the scattering of light from small particles the deflection of any given ray is not a deterministic function of wavelength, but the average intensity of the scattered light does depend on wavelength. Interference gives rise to color through an interaction of light with light, but a change in direction is needed to enable one beam of light to interfere with another.

At the most fundamental level these processes can be understood in terms of electronic excitations in matter. Refraction, for example, results from a change in speed when light passes from one medium into another; the speed in each material is determined by the interaction of the electromagnetic field of the radiation with the electric charges of the electrons. An analysis of this kind is always possible, but it is often too cumbersome to be very informative. What is needed is a "higher level" analysis; it is provided by the methods of geometrical and physical optics.

A light ray is refracted, or bent, by a transparent prism because light is slowed more by traveling in a solid than it is by traveling in air. (The highest possible speed of light, and the speed that relativity theory says can never be exceeded, is attained only in a vacuum; in any material medium the speed is lower.) Why a change in speed should lead to a change in direction can be understood by imagining the light ray as a series of plane wave fronts. When a wave front strikes a transparent surface obliquely, one edge enters and is slowed before the opposite edge reaches the surface; the edge that is slowed first thus falls behind the rest of the wave front, and the ray is deflected toward the per-

perpendicular. On leaving the solid the ray is bent away from the perpendicular.

The magnitude of the change in direction when a light ray is refracted depends on the angle at which it strikes the surface. For any given angle of incidence it also depends on the ratio of the speeds of light in the two materials. It turns out that this ratio, which is called the refractive index, is generally not the same for all wavelengths; that is why refraction can give rise to color. The mechanisms that retard light in a transparent medium have a greater effect on high frequencies and short wavelengths than they do on low frequencies and long wavelengths. As a result violet light is refracted through a larger angle than red light, and a beam of white light on passing through a prism is separated into its component colors.

This dispersion of white light according to wavelength was discovered by Isaac Newton in 1704, or at least it was Newton who first recognized its significance. The dispersion of sunlight by re-

fraction in water droplets or ice crystals is responsible for the colors of a rainbow and of other colored halos occasionally seen around the sun and the moon. Dispersion also causes the "fire," or flashes of color, characteristic of diamond and, to a lesser extent, of other faceted transparent gemstones. Dispersion is not always a welcome phenomenon: in telescopes, cameras and other optical systems it gives rise to chromatic aberration, or the misregistration of images in different colors.

Interference was also first investigated in detail by Newton. It can be observed in a system of waves where two waves combine to yield a single new wave whose amplitude at each point is simply the sum of the amplitudes of the original waves. Thus a monochromatic light wave can be split into two components, which follow different paths and then merge again. Where the two components are in phase, so that the peaks and valleys coincide, the waves reinforce each other and the light is bright. Where

the components are out of phase the waves cancel and appear dark.

Interference is often observed in thin transparent films, where part of the light is reflected by the first surface and part by the second. Whether the beams reinforce or cancel depends on the nature and thickness of the film, on the angle of reflection and on the wavelength of the light. If the film has a uniform thickness, different wavelengths emerge at different angles. If the layer varies in thickness, then at any given viewing angle different colors appear at different positions. Interference in a thin, transparent cuticle gives rise to the colors of some beetles and butterflies. A thin layer of oil on water swirls with color as a result of the same mechanism.

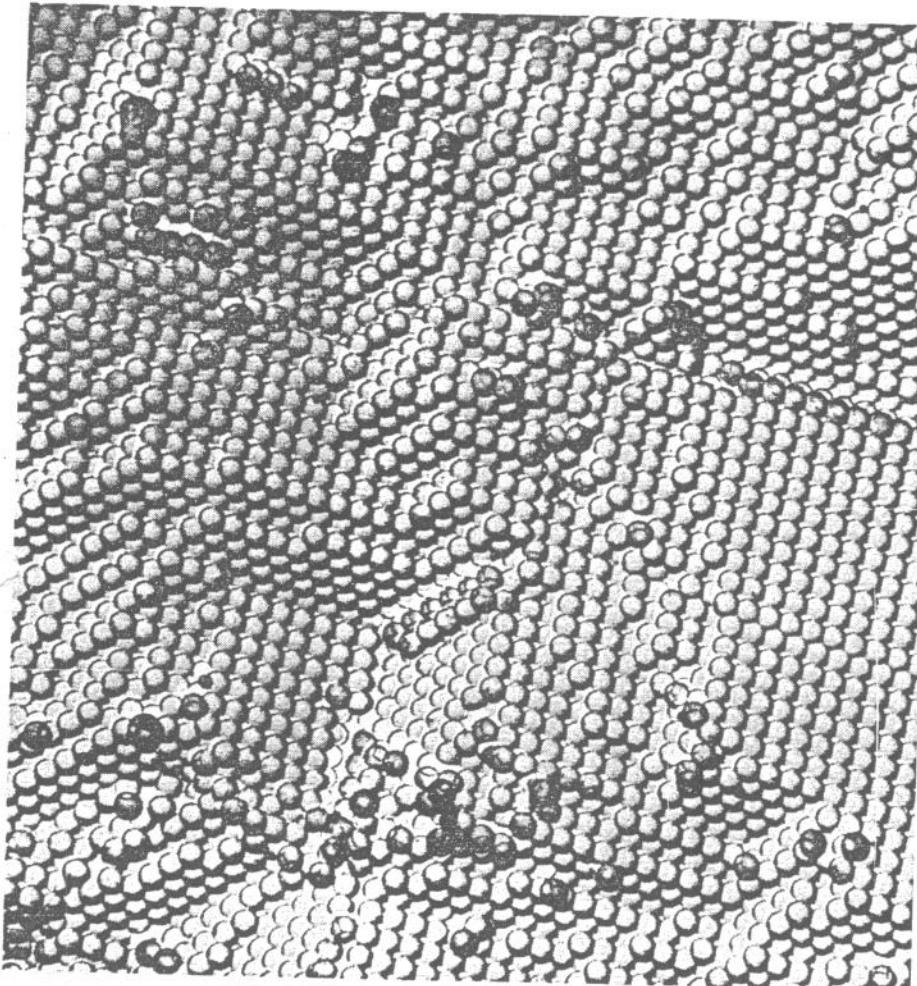
The Diffraction Grating

Diffraction is the bending or spreading of a light wave at the edge of an opaque obstacle. When the obstacle is a macroscopic object, the effect is generally a minor one, but it becomes important when the dimensions of the obstacle are comparable to the wavelength of the light. As might be expected, the magnitude of the effect then becomes strongly dependent on wavelength.

Diffraction is an important cause of color in the scattering of light by small particles. Lord Rayleigh showed that the intensity of the scattered light is inversely proportional to the fourth power of the wavelength; accordingly blue light is scattered about four times as much as red light. The disparity in scattering effectiveness is apparent in the daytime sky. Molecules, dust and fluctuations in density in the atmosphere preferentially scatter blue light, so that the sky appears blue. The direct light of the sun, on the other hand, is depleted of blue and therefore appears reddish, particularly at sunrise and sunset when the light must traverse a greater depth of the atmosphere.

The same process can be demonstrated on a smaller scale by passing white light through diluted unhomogenized milk, where the scattering centers are particles of fat. The fluid appears bluish, whereas the directly transmitted beam is reddish. The blue sheen of moonstone almost certainly comes from a similar process. Scattering from much larger inclusions, which do not alter color, accounts for the patterns in star ruby and star sapphire and in tiger's-eye quartz.

Interference and diffractive scattering act in combination in a final mechanism for generating color: the diffraction grating. Such a grating is an array of many equally spaced lines or points in which the spacing is not too large in comparison with light wavelengths. Light passing through the grating is scattered in all directions at each opening, and so waves from adjacent openings



NATURAL DIFFRACTION GRATING is formed by closely packed spheres in a synthetic opal. The spheres consist of silicon dioxide and a little water, embedded in a transparent matrix that has a similar composition but a slightly different index of refraction. The spheres have a diameter of about 250 nanometers. Dispersion of white light by the three-dimensional diffraction grating gives rise to spectrally pure colors that glint from within an opal. The synthetic opal was made by Pierre Gilson in France, who also made this scanning electron micrograph.

ELECTRONIC TRANSITIONS IN FREE ATOMS AND IONS; VIBRATIONAL TRANSITIONS IN MOLECULES	ELECTRONIC EXCITATIONS	INCANDESCENCE, FLAMES, ARCS, SPARKS, LIGHTNING, GAS DISCHARGES, SOME LASERS.
	VIBRATIONS	BLUE-GREEN TINT OF PURE WATER AND ICE.
CRYSTAL-FIELD COLORS	TRANSITION-METAL COMPOUNDS	TURQUOISE, MOST PIGMENTS, SOME LASERS, SOME PHOSPHORS, SOME FLUORESCENT MATERIALS.
	TRANSITION-METAL IMPURITIES	RUBY, EMERALD, RED SANDSTONE, SOME LASERS, SOME FLUORESCENCE.
	COLOR CENTERS	AMETHYST, SMOKEY QUARTZ, DESERT-AMETHYST GLASS, SOME FLUORESCENCE.
TRANSITIONS BETWEEN MOLECULAR ORBITALS	CHARGE TRANSFER	BLUE SAPPHIRE, MAGNETITE.
	CONJUGATED BONDS	ORGANIC DYES, MOST PLANT AND ANIMAL COLORS, LAPIS LAZULI, FIREFLIES, DYE LASERS, SOME FLUORESCENCE.
TRANSITIONS IN MATERIALS HAVING ENERGY BANDS	METALLIC CONDUCTORS	COPPER, SILVER, GOLD, IRON, BRASS.
	PURE SEMICONDUCTORS	SILICON, GALENA, CINNABAR, DIAMOND.
	DOPED SEMICONDUCTORS	BLUE DIAMOND, YELLOW DIAMOND, LIGHT-EMITTING DIODES, SEMICONDUCTOR LASERS, SOME PHOSPHORS.
GEOMETRICAL AND PHYSICAL OPTICS	DISPERSIVE REFRACTION	THE RAINBOW, "FIRE" IN GEMSTONES, CHROMATIC ABERRATION.
	SCATTERING	BLUE OF THE SKY, RED OF SUNSETS, MOONSTONE, STAR SAPPHIRE.
	INTERFERENCE	OIL FILM ON WATER, LENS COATINGS, SOME INSECT COLORS.
	DIFFRACTION GRATING	OPAL, LIQUID CRYSTALS, SOME INSECT COLORS

CAUSES OF COLOR are classified in 14 categories of five broad types. All but one of the color-causing mechanisms (vibrations of the atoms in molecules) can be traced to changes in the state of the elec-

trons in matter. Electronic transitions are the most important causes of color because the energy needed to excite an electron commonly falls in the range that corresponds to visible wavelengths of light.

interfere. For any one wavelength there are angles at which the interference is constructive; at all other angles the waves cancel one another. The actual value of these angles depends only on the wavelength and on the spacing between lines or points on the grating.

When white light passes through a diffraction grating or is reflected from it, each wavelength is enhanced in a different set of directions. Hence the light in any given direction is spectrally pure and monochromatic, but as the grating is rotated a series of spectra comes into view.

A diffraction grating can be made by inscribing fine rulings on a glass plate, and a few natural systems also have the dimensions and the high degree of order needed for an optical grating. For example, in the materials called liquid crystals molecules are stacked with sufficient regularity to act as a diffraction grating. Because the spacing between molecules depends on the temperature so does the color of the material, and it can be employed as a temperature indicator. Similar ordered arrays of molecules produce diffraction-grating colors in some butterflies and beetles. Spectral

colors arising from diffraction also appear when a distant streetlight is seen through the fabric of an umbrella or when the surface of a phonograph record is viewed at a glancing angle.

The preeminent natural diffraction grating is the opal. In this gemstone spheres made up of silicon dioxide and a little water are closely packed in a three-dimensional array with a spacing of about 250 nanometers. The transparent or translucent matrix that fills the space between spheres has a similar composition but a slightly different index of refraction. When white light enters this three-dimensional diffraction grating, pure colors appear inside the stone; the colors change as the eye or the stone is moved. The mechanism of color production in opal came to be understood only in 1964, when the structure was first resolved with the electron microscope. Synthetic opals were created not long thereafter.

The Visible Spectrum

In this catalogue of colors, which is surely not complete, I have described more than a dozen mechanisms. They

can be put into five broad categories: excitations of free atoms and ions and vibrations in molecules, crystal-field effects, transitions between states of molecular orbitals, transitions in the energy bands of solids and effects interpreted through physical optics. It may seem an extraordinary coincidence that such a diversity of phenomena is encompassed in a band of wavelengths that is not even a full octave wide; it may seem still more remarkable that this narrow band happens to be just the one to which the human eye is sensitive.

Actually it may not be a coincidence at all. So much of interest happens in this narrow region of the electromagnetic spectrum because these are the wavelengths where interactions of light with electrons first become important. Waves of lower energy mainly stimulate the motions of atoms and molecules, and so they are usually sensed as heat. Radiation of higher energy can ionize atoms and permanently damage molecules, so that its effects seem largely destructive. Only in the narrow transition zone between these extremes is the energy of light well tuned to the electronic structure of matter.