

Chapter 2

Radiative quantities

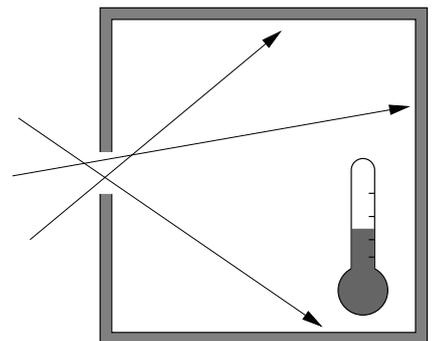
In the radiative transfer approximation we can regard light as photons crisscrossing through space. A photon has an energy and a polarization state, and it has a direction of propagation. At any given time, in a given cubic centimeter of space, one can count the number of photons of a given energy, polarization state and direction that are present. Since usually the number of photons is very large, we typically measure not the photon number but the total energy they represent. The typical quantities we will be concerned with are the bolometric *flux*, *intensity* and *mean intensity*, and their monochromatic versions.

2.1 Radiative flux

We can define the concept of *flux* in the most self-consistent way by defining an experiment that measures this. Let us, as a “gedankenexperiment”, construct a box (“cavity” or “Hohlraum”) with walls that are, on the inside, perfectly black: the walls absorb all photons that impinge on them. On one side we put a small hole with a well-known size. We call this hole the *aperture*. The entire setup is called a *pinhole camera*. Radiation from the outside can enter the aperture and be absorbed by the walls of the cavity. We assume that we know the heat capacity of the walls perfectly and that the walls are perfectly isolated from the exterior. This means that as radiation enters the cavity through the aperture, the cavity starts heating up. A thermometer can measure the increase of the temperature T with time t . If we start from a sufficiently cold cavity, then the thermal radiation from the walls escaping the cavity through the aperture can be ignored compared to the radiation entering the cavity. Under that circumstance, the increase of temperature of the cavity, dT/dt , combined with the known heat capacity of the walls, gives a measurement of the total amount of radiation energy entering the aperture per second. If we divide this by the surface area of the aperture, we get the *radiative flux*: the energy per units of time per unit of surface area. In CGS units this has the dimension of $\text{erg s}^{-1} \text{cm}^{-2}$. The symbol for this is usually F .

The flux that we measure with this device typically depends on the direction in which we point the device. If we point it straight at the sun, we measure a much larger flux than if we point it away from the sun. And if we point it in an angle of 60 degrees from the sun, we measure just $\cos(60^\circ) = 0.5$ times the flux as when we point it directly at the sun. This is indicative of the vectorial nature of flux. The device we constructed measures the component of the flux vector \mathbf{F} that is perpendicular to the surface of the aperture, i.e. it measures $F = \mathbf{n} \cdot \mathbf{F}$, where \mathbf{n} is the normal vector of the aperture, pointing inward into the camera. By measuring the flux in three independent directions we can reconstruct the full flux vector \mathbf{F} .

The flux \mathbf{F} is called the *bolometric flux* because it does not make any distinction between radiation of different wavelengths. It measures the total flux of radiation at



all wavelengths. However, in many applications it is useful to measure flux for specific wavelengths only. One way to do this is to put a color filter in front of the aperture: a piece of glass that only transmits light with wavelengths in some range $\lambda_1 \leq \lambda < \lambda_1 + \Delta\lambda$ and blocks all other radiation. If $\Delta\lambda \ll \lambda_1$ then the temperature increases much slower than before, because much less radiation energy enters the aperture per second. In fact, for sufficiently small $\Delta\lambda$, the radiation energy entering the aperture per second scales linearly with $\Delta\lambda$. The narrower the filter, the less energy enters the aperture. If we do not want our measured quantity to depend on the width of the filter, we can divide the measured flux by $\Delta\lambda$ and we get the *monochromatic flux*. The reason for dividing by $\Delta\lambda$ is similar to the reason why we divided by the surface area of the aperture: we want to measure a physical quantity, not a quantity that depends on the experimental details. The symbol for the monochromatic flux measured by the device is F_λ , and the vectorial flux vector is \mathbf{F}_λ . The monochromatic and bolometric fluxes are related via

$$F = \int_0^\infty F_\lambda d\lambda \quad (2.1)$$

and similar for their vectorial forms. The monochromatic flux F_λ is thus the flux per unit of wavelength. In CGS units this is the flux per cm, and the dimension is $\text{erg s}^{-1} \text{cm}^{-3}$.

Most radiative transfer people, however, prefer a different form of monochromatic flux: F_ν , which is the flux per unit of frequency. In CGS units this is the flux per Hz, and the dimension is $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$. The relation with the bolometric flux is:

$$F = \int_0^\infty F_\nu d\nu \quad (2.2)$$

It is important to realize that

$$F_\nu \neq F_\lambda \quad (2.3)$$

even if, correctly, $\nu = c/\lambda$ (where c is the speed of light). This is because the monochromatic flux is a *distribution function*. It is the flux per unit of wavelength (for F_λ) or per unit of frequency (for F_ν). If we have a filter that transmits light with wavelengths between λ_1 and $\lambda_1 + \Delta\lambda$, then we can alternatively express this as a filter transmitting light between frequencies of $\nu_1 + \Delta\nu$ and ν_1 , where

$$\nu_1 = \frac{c}{\lambda_1} \quad \text{and} \quad \nu_1 + \Delta\nu = \frac{c}{\lambda_1 + \Delta\lambda} \quad (2.4)$$

Note that $\Delta\nu < 0$ for $\Delta\lambda > 0$. For a narrow filter we write $\Delta\lambda \rightarrow d\lambda$ and $\Delta\nu \rightarrow d\nu$ and we can write

$$\frac{d\lambda}{d\nu} = -\frac{c}{\nu^2} = -\frac{\lambda}{\nu} \quad (2.5)$$

With this relation we find a relation between F_λ and F_ν :

$$\lambda F_\lambda = \nu F_\nu \quad (2.6)$$

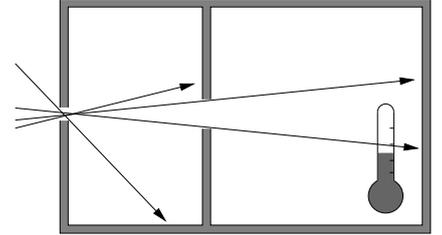
The minus sign disappears because by definition these fluxes are defined to be per *positive* unit of frequency or wavelength respectively.

In this lecture we will always use F_ν .

Note that for historic reasons there are many other units in which astronomers often express a flux. One such unit of flux often used by radio astronomers is Jansky, where $1 \text{ Jy} = 10^{-23} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1}$. Optical astronomers typically use magnitudes, which is a logarithmic scale in which a flux is compared to a standard star (Vega). For a detailed definition of this magnitude system please refer to astronomy introduction books. We will not use these historic systems.

2.2 Radiative intensity

The radiative flux vector contains some directional information, but that information is very incomplete. In radiative transfer theory we usually work with another quantity, called *radiative intensity* (in other fields of physics often called *radiance*). The symbol for this is I for the bolometric case and I_ν for the monochromatic case. Let us again design a simple device to measure it. It is very similar to the flux measuring device, but we add another chamber in front of it with an even smaller aperture than the one in the primary cavity. This new small aperture is now the first aperture, while the aperture of the primary cavity (with the thermometer) is now the second aperture. Any radiation entering the first aperture that does not also pass through the second aperture will be absorbed by the walls of the new chamber in front. We make sure that this heat is quickly removed from the system. Only light that passes through both apertures will enter the main cavity and induce a temperature increase there. This means that the flux that we measure is only the flux from radiation originating from a rather specific direction, namely the direction given by the line-up of the two apertures. The flux should now be defined as the energy per second per unit surface area of the *first* aperture. The second aperture (the entrance of the main cavity with the thermometer) now has the function of selecting radiation from a certain direction and blocking radiation from all other directions.



If the size of the first aperture is small compared to that of the second cavity, then we can in fact easily calculate the total solid angle from which radiation is collected. If the distance between the two apertures is L and the area of the second aperture (the large one) is A , then the solid angle $\Delta\Omega$ from which flux is gathered is:

$$\Delta\Omega = \frac{A}{L^2} \quad (2.7)$$

If $\Delta\Omega \ll 4\pi$ then the measured flux will be proportional to $\Delta\Omega$. So if we divide the measured flux by $\Delta\Omega$ we obtain a “flux per steradian” I :

$$I \equiv \frac{F}{\Delta\Omega} \quad (2.8)$$

This is what we call the *bolometric intensity*. The units of I are $\text{erg s}^{-1} \text{cm}^{-2} \text{ster}^{-1}$. By adding a filter and dividing by $\Delta\nu$ we obtain the *monochromatic intensity* I_ν , which is expressed in units of $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1} \text{ster}^{-1}$.

The intensity is not a vectorial quantity. Instead it is a function of direction. If we define the unit vector \mathbf{n} to point from the small (first) aperture to the second (interior) cavity, we can define the intensity as being a function of \mathbf{n} :

$$I(\mathbf{n}) \quad \text{resp.} \quad I_\nu(\mathbf{n}) \quad (2.9)$$

Note that if our choice of the aperture size is too large, our measurements will be inexact. To get an accurate measurement the first aperture must be much smaller than the second one, and the diameter of the second (interior) aperture must be much smaller than L . The smaller the internal aperture is compared to L the higher the angular resolution of the function $I(\mathbf{n})$. In practice such a pinhole camera is very inefficient since it throws away most of the incident light. The use of lenses will vastly improve the sensitivity of such a device. But the purpose here is not one of efficiency, but just as an illustration of the definitions of flux and intensity.

Aside from the uncertainties introduced by the wave-nature of light, the intensity $I_\nu(\mathbf{n})$ as measured at *one* location \mathbf{x} in space contains almost infinite amount of *angular* information. In fact, our eyes are cameras that measure $I_\nu(\mathbf{n})$ at three frequencies (red, green and blue), and the angular resolution of the images we see is quite staggering. To overcome the limitation of the wave-nature of light we would have to increase our camera: larger telescopes have larger angular resolution. So from a strictly physical perspective the intensity function $I_\nu(\mathbf{n})$ is a function with limited angular resolution,

depending on the size of the camera. In the radiative transfer approximation, however, we assume it to have infinite angular resolution and to be well-defined at each point \mathbf{x} in space:

$$I_\nu(\mathbf{x}, \mathbf{n}) \quad (2.10)$$

It is thus a 6-dimensional function: It depends on frequency, on three spatial dimensions and on 2 angular directions (note that while \mathbf{n} has three components, only two are independent).

The intensity function, even when considered only at one single point in space, contains an incredible amount of information. If one were to know the function $I_\nu(\mathbf{x}, \mathbf{n})$ at the position of the Earth to infinite precision, it contains all the astronomical observations we will ever be able to do (aside from time-dependent phenomena, of course). One could say that large astronomical all-sky surveys such as the 2MASS infrared survey measure this function for one (or a few) values of ν and one value of \mathbf{x} (the position of the Earth). In many radiative transfer problems, however, we must deal with the intensity at many locations \mathbf{x} and many frequencies ν simultaneously. This may give a feeling for the complexity of the problem. Fortunately, for most of these problems we do not have to calculate the angular dependence to such high resolution. This keeps these problems tractable.

2.3 Angular coordinates

So far we used a unit vector \mathbf{n} to denote direction. We thus write the direction-dependency of intensity as $I(\mathbf{n})$. However, this can sometimes be a bit too abstract. The vector \mathbf{n} should represent the two angular coordinates, but in fact it has 3 components. Of course, the constraint that $|\mathbf{n}| = 1$ removes the 3rd component, but it remains somewhat implicit.

A more concrete way of writing direction is by using angles θ and ϕ . However, this requires us to choose a reference direction. Let us choose the z -axis for that. This choice breaks the symmetry of the problem, which is somewhat unelegant. But it allows us to be more explicit in our writing of direction. We choose the angle θ to be the angle between the \mathbf{n} vector and the \mathbf{e}_z basis vector:

$$\cos \theta = \mathbf{n} \cdot \mathbf{e}_z \quad (2.11)$$

We define the angle ϕ as the angle between the projected vector

$$\tilde{\mathbf{n}} = \mathbf{n} - (\mathbf{n} \cdot \mathbf{e}_z)\mathbf{e}_z \quad (2.12)$$

and the \mathbf{e}_x basis vector

$$\cos \phi = \frac{\tilde{\mathbf{n}} \cdot \mathbf{e}_x}{|\tilde{\mathbf{n}}|} \quad (2.13)$$

where $0 < \phi < \pi$ means that $\tilde{\mathbf{n}} \cdot \mathbf{e}_y > 0$.

Reversely, given θ and ϕ we can construct the components of \mathbf{n} as

$$n_x = \sin \theta \cos \phi \quad (2.14)$$

$$n_y = \sin \theta \sin \phi \quad (2.15)$$

$$n_z = \cos \theta \quad (2.16)$$

Using θ and ϕ we can now express the intensity as

$$I_\nu(x, y, z, \theta, \phi) \quad (2.17)$$

where x , y and z are the three spatial coordinates. This functional form is explicit, but one must always keep in mind that the θ , ϕ coordinates are singular at $\theta = 0$ and $\theta = \pi$.

For reasons that will become apparent soon, it is often convenient to use, instead of θ :

$$\mu \equiv \cos \theta \quad (2.18)$$

We then get

$$n_x = \sqrt{1 - \mu^2} \cos \phi \quad (2.19)$$

$$n_y = \sqrt{1 - \mu^2} \sin \phi \quad (2.20)$$

$$n_z = \mu \quad (2.21)$$

and we can express the intensity as

$$I_\nu(x, y, z, \mu, \phi) \quad (2.22)$$

In the rest of this lecture we use \mathbf{n} , (θ, ϕ) and (μ, ϕ) interchangeably, depending on what is more convenient for the problem at hand. It will be clear from the context which of these three notations we use.

2.4 Intensity is constant along a ray - in vacuum

The intensity I (or I_ν) has an amazing, and extremely useful property: If we are following a ray of light in vacuum, then the intensity in the direction of the ray is constant along that ray! Mathematically one can express this as

$$\mathbf{n} \cdot \nabla I_\nu(\mathbf{x}, \mathbf{n}) = 0 \quad (2.23)$$

where ∇ is the gradient operator and thus $\mathbf{n} \cdot \nabla$ is the derivative taken in the direction of \mathbf{n} .

This means that if we measure the intensity I_ν of the Sun from Earth, and we redo the measurement on Mercury, we get the same answer. How can this be? It seems to be in conflict with the known fact that the flux F_ν from the sun scales with $1/r^2$, i.e. inversely proportional to the distance squared. The key to this apparent paradox is that if we make a photo of the Sun from Earth and then, with the same camera, a photo of the Sun from Mercury, then on the latter picture the sun looks larger. It covers a larger $\Delta\Omega$ on the photo, i.e. it spreads its radiation over a larger number of pixels of your image. Since the angular scale of the Sun scales as $\Delta\Omega \propto 1/r^2$ and so does the flux F_ν , their ratio stays constant with r . In other words: the intensity is independent of r .

In fact, if you make a photo with a digital camera, or if you look at something with your eyes, then what your image records is the intensity in each direction.

So how is this consistent with the fact that distant stars look dim? Indeed, even a very distant star has the same intensity as the sun (if it has the same temperature). However, in order to measure the intensity of the star, your telescope must have an angular resolution that is sufficiently large to be able to resolve the surface of that star. For most stars that is impossible. The stellar flux will be spread over a pixel that is much larger than the projected size of the star on the image. This de-facto dilutes the intensity, since the measured flux will then not be divided by the solid angle of the star, but by the solid angle of the pixel. The intensity of the star therefore might look weaker than the Sun, but that is merely an artifact of the finite angular resolution of your telescope.

Eq. (2.23) can also be written as a differential equation along a straight line tangent to the direction vector \mathbf{n} . Pick any straight ray you wish, and define a coordinate s along that ray that denotes a distance along that ray such that the ray is defined as

$$\mathbf{x}(s) = \mathbf{x}_0 + s\mathbf{n} \quad (2.24)$$

Then Eq. (2.23) can be cast into the form

$$\frac{dI_\nu(\mathbf{n})}{ds} = 0 \quad (2.25)$$

or in other words, that $I_\nu(\mathbf{n})$ is constant along that ray. This does not mean that $I_\nu(\mathbf{n})$ should be constant throughout space, since (a) it can be different along different parallel rays and (b) it can be different for different directions \mathbf{n} .

The constancy of intensity along a ray plays a fundamental role in radiative transfer theory, and Eqs.(2.23 and 2.25) stand at the basis of the radiative transfer equation which we shall discuss in Section 3.2.

2.4.1 Mirrors and lenses: Intensity still stays constant

This constancy of intensity remains true even if the ray gets reflected off a mirror, even if the mirror is curved. If you look at yourself in a concave mirror (e.g. a spoon) you see yourself bigger, but your skin does not look brighter or dimmer. The same is true if you use a magnifying glass: things are magnified, but remain as bright as before. This all does not depend on details of the optical setup. The reason is simple: without this constancy of intensity rule we would violate the rules of thermodynamics.

2.5 Moments of intensity

As mentioned at the end of Section 2.2, the richness of angular information packed in the function $I_\nu(\mathbf{x}, \mathbf{n})$ can make the problem of radiative transfer very hard to solve. But it was also mentioned that often one can afford a much lower angular resolution and thus make the problem more feasible. An elegant way of lowering angular resolution in a controlled way is to expand the angular dependency of I_ν into spherical harmonics. The lowest order components represent the lowest resolution information while successively higher order components contain information at successively higher angular resolution. One of the disadvantages of a spherical harmonics expansion is that one must specify an axis of reference, which breaks any possible rotational symmetries of the problem.

An equivalent method of expansion, and one that does not introduce a preferential reference direction, is the expansion into tensor moments. A complete and mathematically rigorous exposé of the expansion into tensor moments is given by Thorne (1981, MNRAS, 194, 439).

For our purposes, however, we will remain far less complete. We will define only the zeroth, first and second tensor moment of the radiation field (where, for notational convenience, we will omit the dependence on \mathbf{x}):

$$J_\nu = \frac{1}{4\pi} \oint I_\nu(\mathbf{n}) d\Omega \quad (2.26)$$

$$\mathbf{H}_\nu = \frac{1}{4\pi} \oint I_\nu(\mathbf{n}) \mathbf{n} d\Omega \quad (2.27)$$

$$\mathcal{K}_\nu = \frac{1}{4\pi} \oint I_\nu(\mathbf{n}) \mathbf{n} \mathbf{n} d\Omega \quad (2.28)$$

These are integrals over all directions with $d\Omega$ the solid angle and \mathbf{n} the direction.

The zeroth moment J_ν is called the *mean intensity* and is indeed the angular average of $I_\nu(\mathbf{n})$. If we are in a homogeneous and isotropic radiation field, then $J_\nu = I_\nu$.

The first moment \mathbf{H}_ν is a vectorial quantity that is, in fact, identical to the flux, apart from a factor $1/4\pi$:

$$\mathbf{F}_\nu = \oint I_\nu(\mathbf{n}) \mathbf{n} d\Omega = 4\pi \mathbf{H}_\nu, \quad (2.29)$$

Often \mathbf{F}_ν and \mathbf{H}_ν are used interchangeably, whichever turns out to be convenient in the particular context. This quantity denotes the average flow of radiative energy. For a homogeneous and isotropic radiation field $\mathbf{H}_\nu = 0$.

The second moment \mathcal{K}_ν is a symmetric tensor of rank 2. It is the quantity that can be interpreted as being responsible for radiation pressure, but it also has its uses when radiation pressure is weak. For a homogeneous and isotropic radiation field $\mathcal{K}_\nu = \frac{1}{3}\infty J_\nu$, where ∞ is the unit rank-2 tensor.

We could define ever higher rank moment tensors, and for a complete description of the radiation field we would indeed need to use an infinite series of moments. However, as we shall see later, for most purposes these first three moments are sufficient. We will use the moment formalism later to derive the equations for radiative diffusion (Section 4.5).

With the angular coordinates defined in Section 2.3 we can write the integrals of Eqs. (2.26, 2.27, 2.28) in a more explicit form:

$$J_\nu = \frac{1}{4\pi} \int_{-1}^{+1} d\mu \int_0^{2\pi} d\phi I_\nu(\mu, \phi) \quad (2.30)$$

$$H_\nu^i = \frac{1}{4\pi} \int_{-1}^{+1} d\mu \int_0^{2\pi} d\phi I_\nu(\mu, \phi) n^i \quad (2.31)$$

$$K_\nu^{ij} = \frac{1}{4\pi} \int_{-1}^{+1} d\mu \int_0^{2\pi} d\phi I_\nu(\mu, \phi) n^i n^j \quad (2.32)$$

where $i = 1, 2, 3$ (denoting the directions x, y and z respectively) and n^i are the components of the \mathbf{n} vector given by Eqs. (2.19, 2.20, 2.21).

2.6 Thermal radiation: The Planck function

Now suppose we take our cavity again, but close the aperture. We will soon obtain a thermodynamic equilibrium at some temperature T inside the cavity. This also means that the cavity will be filled with thermal radiation of temperature T . The walls will continuously emit and absorb such thermal radiation. The photon quantum states inside the cavity will then have an occupation number according to Bose-Einstein statistics:

$$N = \frac{1}{e^{h\nu/kT} - 1} \quad (2.33)$$

where $\epsilon = h\nu$ is the energy of that quantum state. Since the density of quantum states (per volume per frequency) is $\rho_s = 4\pi g\nu^2/c^3$ (with $g = 2$ for photons because each photon of given energy $h\nu$ can have two independent polarization states), we see that the equilibrium energy density for light $U(\nu)$ in $\text{erg cm}^{-3} \text{ Hz}^{-1}$ is

$$U(\nu) = \frac{4\pi g h\nu^3/c^3}{e^{h\nu/kT} - 1} = \frac{8\pi h\nu^3/c^3}{e^{h\nu/kT} - 1} \quad (2.34)$$

Now, per steradian (dividing by 4π) and passing through a surface of 1 cm^2 per second (multiplying with c) this gives the *Planck function*:

$$B_\nu(T) = \frac{2h\nu^3/c^2}{e^{h\nu/kT} - 1} \quad (2.35)$$

Inside the thermal cavity the intensity of the radiation in any direction is then

$$I_\nu(\mathbf{n}) = B_\nu(T) \quad (2.36)$$

This radiation field inside the thermal cavity is called *blackbody radiation*. This also means that the thermal emission from a thermal surface at temperature T is such that the intensity seen from that surface is $I_\nu = B_\nu(T)$, independent of the direction from which you look at that surface.

This knowledge can now also be used outside the context of a closed cavity. If we have a perfectly thermally emitting black surface anywhere, we can be assured that the intensity I_ν seen from that surface *from any direction* equals $B_\nu(T)$, where T is the temperature of that surface. This is what is meant with a *blackbody emitter*.

We can calculate the flux emitted by a thermal blackbody surface. Let us put the thermal surface in the (x, y) -plane, so that the normal of that surface points upward along the z -axis. We can then use Eq. (2.31) to compute the z -component of \mathbf{H}_ν :

$$H_\nu = \frac{1}{2} \int_0^{+1} B_\nu(T) \mu d\mu = \frac{1}{4} B_\nu(T) \quad (2.37)$$

Note that the μ integration domain goes here from 0 to 1, instead of -1 to 1, because we study here only the radiation emitted from the surface, which has $\mu > 0$. The flux F_ν is then:

$$F_\nu = \pi B_\nu(T) \quad (2.38)$$

When we integrate this over all frequencies, we get the bolometric flux:

$$F = \int_0^\infty F_\nu d\nu = \pi \int_0^\infty B_\nu(T) d\nu = \sigma_{\text{SB}} T^4 \quad (2.39)$$

where $\sigma_{\text{SB}} = 5.67 \times 10^{-5} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ K}^{-4}$ is the Stefan-Boltzmann constant. Note, by the way, that this gives the following expression for the bolometric Planck function:

$$B(T) \equiv \int_0^\infty B_\nu(T) d\nu = \frac{\sigma_{\text{SB}}}{\pi} T^4 \quad (2.40)$$

Typically, a real surface is not a black body. It has a certain thermal efficiency ϵ_ν such that the thermal flux from that surface is

$$F_\nu = \epsilon_\nu \pi B_\nu(T) \quad (2.41)$$

As a consequence, if the surface is opaque, it must have a reflectivity $\eta_\nu = 1 - \epsilon_\nu$, which is called the *albedo*. Objects in every day life have an albedo η_ν that varies with frequency ν , which gives these objects color.