Part 1 Ray Optics

In this part, we treat light beams as rays that propagate along straight lines, except at interfaces between dissimilar materials, where the rays may be bent or refracted. This approach, which had been assumed to be completely accurate before the discovery of the wave nature of light, leads to a great many useful results regarding lens optics and optical instruments.

1.1 Refraction and Reflection

1.1.1 Refraction

When a light ray strikes a smooth interface between two transparent media at an angle, it is refracted. Each medium may be characterized by an index of refraction *n*, which is a useful parameter for describing the sharpness of the refraction at the interface. The index of refraction of air (more precisely, of free space) is arbitrarily taken to be one, *n* is most conveniently regarded as a parameter whose value is determined by experiment. We know now that the physical significance of *n* is that the ratio of the velocity of light in vacuo to that in the medium.

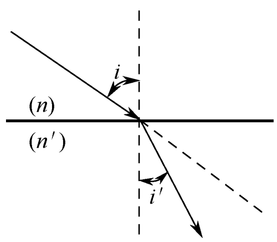

Fig.1.1 Refraction at an interface

Suppose that the ray is incident on the interface, as shown in Fig.1.1. It is refracted in such a way that

$$n \sin i = n' \sin i' \tag{1.1}$$

no matter what the inclination of the incident ray to the surface, *n* is the index of refraction of the first medium, n' that of the second. The angle of incidence *i* is the angle between the incident ray and the normal to the surface; the angle of refraction i' is the angle between the refracted ray and the normal.

1.1.2 Index of Refraction

Most common optical materials are transparent in the visible region of the spectrum, whose wavelength ranges from 400 to 700nm. They exhibit strong absorption at shorter wavelengths, usually 200nm and below.

The refractive index of a given material is not independent of wavelength, but generally increases slightly with decreasing wavelength (Near the absorption edge at 200 nm, the index of glass increases sharply). This phenomenon is known as dispersion. Dispersion can be used to display a spectrum with a prism; it also gives rise to unwanted variations of lens properties with wavelength. Table 1.1 gives typical index of refraction of several materials.

Tab.1.1 Index of refraction of several materials

Material	Index of refraction	Material	Index of refraction
air	1.0003	sodium chloride	1.54
water	1.33	light flint glass	1.57
magnesium fluoride	1.38	Sapphire	1.77
vitreous silica	1.46	extra-dense flint glass	1.73
Pyrex glass	1.47	carbon disulfide	1.62
Methanol	1.33	zinc sulfide (thin film)	2.3
xylene	1.50	medium flint glass	1.63
ethanol	1.36	titanium dioxide (thin film)	2.4～2.9
crown glass	1.52	heaviest flint glass	1.89
benzene	1.50	Canada balsam (center)	1.53

Optical glasses are generally specified both by index n and by a quantity known as dispersion ν,

$$\nu = \frac{n_F - n_C}{n_D - 1} \tag{1.2}$$

The subscripts F, D and C refer to the indexes at certain short, middle and long wavelengths (blue, yellow, red).

1.1.3 Reflection

Certain highly polished metal surfaces and other interfaces may reflect all or nearly all of the light falling on the surface. In addition, ordinary, transparent glasses reflect a few percent of the incident light and transmit the rest.

The angle of incidence is i and the angle of reflection i'. Experiment shows that the angles of incidence and reflection are equal, except in a very few peculiar cases, as shown in Fig.1.2.

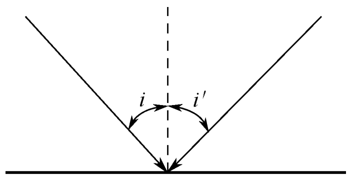

Fig.1.2 Reflection at an interface

We shall later adopt the convention that i is positive; that is, if the acute angle opens counterclockwise from the normal to the ray, i is positive. The sign of i' is clearly opposite to that of i. We therefore write the law of reflection as

$$i' = -i \tag{1.3}$$

1.1.4 Total Internal Reflection

Here we consider a ray that strikes an interface from the high-index side, say, from glass to air (not air to glass). This is known as internal reflection. The law of refraction shows that the incident ray is in this case bent away from the normal when it crosses the interface, as shown in Fig.1.3.

Thus, there will be some angle of incidence for which the refracted ray will travel just parallel to the interface. In this case, $i' = 90°$, so the law of refraction becomes

$$n \sin i_c = n' \sin 90^\circ \tag{1.4}$$

where i_c is known as the critical angle. Since $\sin 90^\circ = 1$,

$$\sin i_c = n' / n \tag{1.5}$$

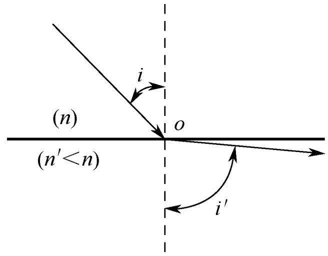

Fig.1.3 Refraction near the critical angle

If i exceeds i_c, then $n\sin i > n'$, and the law of refraction demands that $\sin i'$ exceed 1. Because this is impossible, we can conclude only that there can be no refracted ray in such cases. The light cannot simply vanish, so we are not surprised that it must be wholly reflected; this is indeed the case. The phenomenon is known as total internal reflection; it occurs whenever

$$i > \arcsin(n' / n) \tag{1.6}$$

The reflected light, of course, obeys the law of reflection.

For a typical glass-air interface, n=1.5, the critical angle is about 42°. Glass prisms that exhibit total reflection are therefore commonly used as mirrors with angles of incidence of about 45°.

1.1.5 Reflecting Prisms

There are different types of reflecting prism. The most common are prisms whose cross sections are right isosceles triangles. One advantage of a prism over a metal-coated mirror is that its reflectance is nearly 100% if the surfaces normal to the light are antireflection coated. Further, the prism's properties do not change as the prism ages, whereas metallic mirrors are subject to oxidation and are relatively easy to scratch. A glass prism is sufficiently durable that it can withstand all but the most intense laser beams. Fig.1.4 shows a prism being used in place of a plane mirror.

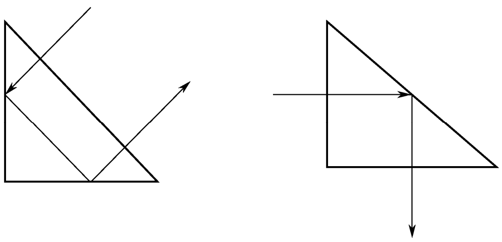
Fig.1.4 Reflecting prisms

In imaging-forming systems, these prisms must be used in collimated light beams to avoid introducing defects into the optical image.

1.2 Imaging

1.2.1 Spherical Surfaces

Because a simple lens consists of a piece of glass with, in general, two spherical surfaces, we will find it necessary to examine some of the properties of a single, spherical refracting surface. We will for brevity call such a surface, as shown in Fig.1.5, a "len". Two of these form a lens. To avoid confusion, we will always place "len" in quotes.

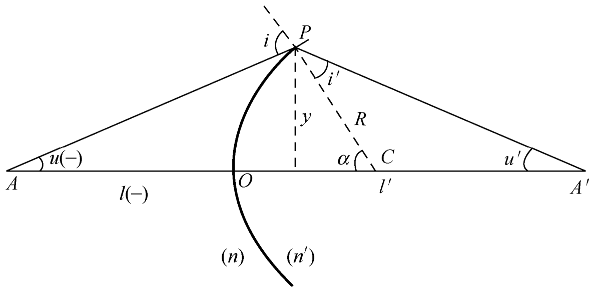

Fig.1.5 Spherical refracting surface

We are interested in the imaging property of the 'len". We consider a bright point A and define the axis along the line *AC*, where *C* is the center of the spherical surface. We examine a particular ray *AP* that strikes the "len" at *P*. We shall be interested in the point A' where this ray intersects the axis.

Before proceeding any further, we must adopt a sign convention. The choice of convention is, of course, arbitrary, but once we choose a convention, we shall have to stick with it. The convention we adopt appears, at first, quite complicated. We choose it at least in part because it is universally applicable; with it we will not need to derive a special convention for spherical mirrors.

To begin, imagine a set of Cartesian coordinate axes centered at *O*. Distances are measured from *O*. Distances measured from *O* to the right are positive; those measured from *O* to the left are negative. Thus, for example, OA' and *OC* are positive, whereas *OA* is negative. Similarly, distances measured above the axis are positive; those below are negative. This is our first sign convention.

We now adopt a convention for the signs of angles such as *OAP* or $OA'P$. We determine their signs by trigonometry. For example, the tangent of angle *OAP* is approximately

$$\tan OAP \approx y/OA \tag{1.7}$$

where y is the distance indicated between *P* and the axis. Our previous convention shows that y is positive, and *OA*, negative. Thus, tan *OAP* is negative and so is *OAP* itself. Similarly, $OA'P$ and *OCP* are positive.

This is our second sign convention. An equivalent statement is that angle $OA'P$ (for example) is positive if it opens clockwise from the axis, or negative otherwise. It is probably simplest, however, merely to remember that angle *OAP* is negative as drawn in Fig.1.5.

Finally, we deal with angles of incidence and refraction, such as angle CPA'. It is most convenient to define CPA' to be positive as shown in Fig.1.5. The angle of incidence or refraction is positive if it opens counterclockwise from the normal (which is, in this case, the radius of the spherical surface）.

Unfortunately, when the last convention is expressed in this way, the statement differs from that which refers to angles (such as *OAP*) formed by a ray crossing the axis. It is best to learn the sign convention by remembering the signs of all of the important angles in Fig.1.5. Only angle *OAP* is negative.

Let us now assign symbols to the more important quantities in Fig.1.5. The point A' is located a distance l' to the right of *O*, and the ray intersects the axis at A' with angle u'. The radius *R* through the point *P* makes angle α with the axis. The angles of incidence and refraction are i and i', respectively.

We must be careful of the signs of *OA* and angle *OAP*, both of which are negative according to our sign convention. This is indicated in Fig.1.5 with parenthetical minus signs. We shall later find it necessary, after a derivation based on geometry alone, to go through our formulas and change the signs of all quantities that are algebraically negative. This is so because our sign convention is not used in ordinary geometry. To make our formulas both algebraically and numerically correct, we must introduce our sign convention, which we do as indicated, by changing signs appropriately.

1.2.2 Object-Image Relationship

We now attempt to find a relationship between the quantities l and l' for a given geometry. First, we relate angle u and i to angle α. The three angles in triangle PAC are u, α and $\pi - i$. Because the sum of these angles must be π, we have

$$u + \alpha + (\pi - i) = \pi \tag{1.8}$$

or

$$i = \alpha + u \tag{1.9}$$

Similarly

$$i' = \alpha - u' \tag{1.10}$$

At this point, it is convenient to make the paraxial approximation, namely, the approximation that the ray AP remains sufficiently close to the axis that angles u, u', i and i' are so small that their sines or tangents can be replaced by their arguments; that is

$$\sin\theta = \tan\theta = \theta \tag{1.11}$$

where θ is measured in radians.

It is difficult to draw rays that nearly coincide with the axis, so we redraw Fig.1.5 by expanding the vertical axis a great amount, leaving the horizontal axis intact. The vertical axis has been stretched so much that the surface looks like a plane. In addition, because only one axis has been expanded, all angles are greatly distorted and can be discussed only in terms of their tangents. Thus, for example,

$$u = y / l \tag{1.12}$$

and

$$u' = y / l' \tag{1.13}$$

in paraxial approximation. Note also that large angles are distorted. Although the radius is normal to the surface, it does not look normal in the paraxial approximation.

To return to the problem at hand, the law of refraction is

$$ni = n'i' \tag{1.14}$$

in paraxial approximation, from which we write

$$n(\alpha + u) = n'(\alpha - u') \tag{1.15}$$

Because $OC=R$, we write α as

$$\alpha = y / R \tag{1.16}$$

The last equation therefore becomes

$$n\left(\frac{y}{R} + \frac{y}{l}\right) = n'\left(\frac{y}{R} - \frac{y}{l'}\right) \tag{1.17}$$

A factor of y is common to every term and therefore cancels. We rewrite this relation as

$$\frac{n'}{l'} + \frac{n}{l} = \frac{n' - n}{R} \tag{1.18}$$

At this point, we have made no mention of the sign convention. We derived the proceeding equation on the basis of geometry alone. According to our sign convention, all of the terms in the equation are positive, except l, which is negative. To make the equation algebraically correct, we must, therefore, change the sign of the term containing l. This change alters the equation to

$$\frac{n'}{l'}-\frac{n}{l}=\frac{n'-n}{R} \tag{1.19}$$

which we refer to as the “len” equation.

There is no dependence on y in the “len” equation. Thus, in paraxial approximation, every ray leaving A (and striking the surface) crosses the axis at A'. We therefore refer to A' as the image of A. A and A' are called conjugate points, and the object distance l and image distance l' are called conjugates.

Had we not made the paraxial approximation, the y dependence of the image point would not have vanished. Rays that struck the lens at large values of y would not cross the axis precisely at A'. The dependence on y is relatively small, so we would still refer to A' as the image point. We say that the image suffers from aberrations if all of the geometrical rays do not cross the axis within a specified distance of A'.

1.2.3 Use of the Sign Conventions

A word of warning with regard to the signs in algebraic expression: Because of the sign convention adopted here, derivations based solely on geometry will not necessarily result in the correct sign for a given term. There are two ways to correct this defect. The first, to carry a minus sign before the symbol of each negative quantity, is too cumbersome and confusing for general use. Thus, we adopt the second, which is to go through the final formula and change the sign of each negative quantity. This procedure has already been adopted in connection with the “len” equation and is necessary, as noted, to make the formula algebraically correct. It is important, though, not to change the signs until the final step, lest some signs be altered twice.

1.2.4 Lens Equation

A thin lens consists merely of two successive spherical refracting surfaces with a very small separation between them. Fig.1.6 shows a thin lens in air. The index of the lens is n. The two refracting surfaces have radii R_1 and R_2, both of which are drawn positive.

We can derive an equation that relates the object distance l and the image distance l' by considering the behavior of the two surfaces separately. The first surface alone would project an image of point A to a point A_1'. If A_1' is located at a distance l_1' to the right of the first surface, the “len” equation shows that, in paraxial approximation,

$$\frac{n}{l_1'}-\frac{1}{l}=\frac{n-1}{R_1} \tag{1.20}$$

because n is the index of the glass (second medium) and 1, the index of the air.

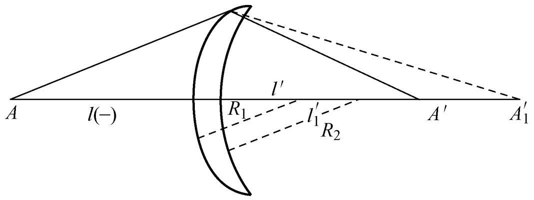

Fig.1.6 Thin lens

The ray does not ever reach A_1', because it is intercepted by the second surface. The second surface, however, behaves as if an object were located at A_1'. The object distance is l_1', if we neglect the thickness of the lens. In applying the "len" equation to the second surface, we must realize that the ray travels across the interface from glass to air. Thus *n* is the index of the first medium and 1, that of the second. The final image point A' is also the image projected by the lens as a whole. If we call the corresponding image distance l', then the "len" equation yields

$$\frac{1}{l'}-\frac{n}{l_1'}=\frac{1-n}{R_2} \tag{1.21}$$

for the second surface.

If we add the last two equations algebraically, we find that

$$\frac{1}{l'}-\frac{1}{l}=(n-1)\left(\frac{1}{R_1}-\frac{1}{R_2}\right) \tag{1.22}$$

which is known as the lens-maker's formula. The lens-maker's formula was derived from the "len" equation by algebra alone. There are no signs to change because that step was included in the derivation of the "len" equation.

We may define a quantity f' whose reciprocal is equal to the right-hand side of the lens-maker's formula,

$$\frac{1}{f'}=(n-1)\left(\frac{1}{R_1}-\frac{1}{R_2}\right) \tag{1.23}$$

The lens-maker's formula may then be written as

$$\frac{1}{l'}-\frac{1}{l}=\frac{1}{f'} \tag{1.24}$$

where f' is the focal length of the lens. We call this equation the lens equation.

We may see the significance of f' in the following way. If the object is infinitely distant from the lens, then $l=-\infty$. The lens equation then shows that the image distance is equal to f'. If the object is located along the axis of the lens, the image also falls on the axis. We call the image point in this case the secondary focal point F'. Note that any ray that travels parallel to the axis is directed by the lens through F', an observation that we will later find particularly useful.

We define the primary focal point *F* in a similar way. The primary focal length *f* is the object distance for which $l'=\infty$. Thus, the lens equation shows that

$$f'=-f \tag{1.25}$$

the primary and secondary focal lengths have equal magnitudes. Any ray that passes through *F* will be directed by the lens parallel to the axis.

Finally, we note that, in the general case, a lens may have different media on opposite sides. In this case, the lens equation may be shown to be

$$\frac{n'}{l'}-\frac{n}{l}=\frac{n'}{f'}=-\frac{n}{f} \tag{1.26}$$

where n and n' are the indices in the first and second media, respectively. The primary and secondary focal lengths are not equal, but are related by

$$\frac{f'}{f}=-\frac{n'}{n} \tag{1.27}$$

1.2.5 Classification of Lenses and Images

A positive lens is a lens that will cause a bundle of parallel rays to converge to a point. Its secondary focal point lies to the right of the lens and f' is therefore positive. It may be regarded as a lens that is capable of projecting an image of a relatively distant object on a screen. An image that can be projected on a screen is called a real image. In general, a positive lens projects a real, inverted image of any object located to the left of its primary focal point *F*. When an object is located at *F*, the image is projected to ∞. The lens is not strong enough to project an image when the object is inside *F*. In that case, an erect image appears to lie behind the lens and is known as a virtual image.

A positive lens need not have two convex surfaces. It may have the meniscus shape of Fig.1.6. If the lens is thickest in the middle, the lens-maker's formula will show it to be a positive lens.

A negative lens has its secondary focal point located to the left. Its secondary focal length f' is negative, and it cannot project a real image of a real object. Rather, it displays an erect, virtual image of such an object. In only one instance can a negative lens display a real image. This is the case when a positive lens projects a real image that is intercepted by a negative lens located to the left of the image plane. Because the rays are cut off by the negative lens, the real image never appears, but behaves as a virtual object projected by the negative lens.

Like a positive lens, a negative lens need not be concave on both surfaces, but may be a meniscus. If the lens is thinnest in the center, f' will prove to be negative and the lens, also negative.

1.2.6 Spherical Mirrors

Our formalism allows mirror optics to be developed as a special case of lens optics. We notice first that the law of reflection $i'=-i$ can also be written

$$(-1)\sin i' = 1\sin i \tag{1.28}$$

which is precisely analogous to the law of refraction, with $n'=-1$. We may therefore regard a mirror as a single refracting surface, across which the index changes from +1 to −1. It is left as a problem to apply the "len" equation to this case. We find that the focal length of a mirror is

$$f' = R/2 \tag{1.29}$$

where *R* is the radius of curvature. In addition, the focal points *F* and F' coincide. The formula that relates the conjugates for a curved-mirror system is

$$\frac{1}{l'}+\frac{1}{l}=\frac{2}{R} \tag{1.30}$$

Mirrors are usually classified as concave and convex. A concave mirror usually projects a real, inverted image, whereas a convex mirror forms an erect, virtual image.

1.2.7 Aberrations

The aberrations of simple, single-element lenses can be quite severe when the lens is comparatively large (with respect to image or object distance) or when the object is located far from the lens axis. When a simple lens is incapable of performing a certain task, it will be necessary to employ a lens, such as a camera lens, whose aberrations have been largely corrected. For specially demanding functions, special lenses may have to be designed and built.

All real lenses made from spherical surfaces may display spherical aberration. Additionally, if the object point is distant from the axis of the lens, or off-axis, the image may display other aberrations, such as astigmatism, coma, distortion, and field curvature. Furthermore, the index of refraction of the lens is a function of wavelength, so its focal length varies slightly with wavelength; the resulting aberration is called chromatic aberration.

Spherical aberration appears both on the axis and off the axis, and does not depend on the distance off-axis. Astigmatism occurs because an off-axis bundle of rays strikes the lens asymmetrically. This asymmetry causes a pair of line images to appear: one behind the plane of best focus and the other in front of it. Coma gives rise to a cometlike image; the head of the comet is the paraxial image point, and the aberration manifests itself as the tail. The tail points away from the axis of the lens and is 3 times longer than its wide. The length of the comatic image, from the paraxial image point to the end of the tail, increases in proportion to the square of the lens diameter and to the distance of the image point from the axis of the lens. The image projected by a lens does not truly lie on a plane but rather on a curved surface, even if other aberrations are zero. This aberration is called field curvature. If the magnification is function of the distance of an image point from the axis, then the image will not be rectilinear. The resulting aberration is called distortion.

Aberrations may be reduced by adjust the radii of the curvature of lens elements so that, for example, angles of incidence are minimized; this process is sometimes called bending the lens. Astigmatism, however, is only weakly influenced by bending the elements. Similarly, one aberration can sometimes be balanced against another. For example, spherical aberration can be partially compensated by moving the image plane from the paraxial image plane to the waist, that is, by compensating spherical aberration by defocusing. Similarly, coma, distortion, and astigmatism can be reduced by adjusting the axis position of the aperture stop.

Words and Expressions

a bundle of	一束
aberration	像差
acute	（尖）锐的，锐角的
algebra /algebraically	代数/用代数的方法
all but	几乎

antireflection	减反射，增透
applicable	可适用的，能应用的，合适的，适当的
approach	方法，路径
arbitrarily	人为地
astigmatism	像散
asymmetry	非对称性
be analogous to	与……类似
be subject to	常遭受
Cartesian coordinate	笛卡儿坐标系
capable	能干的，能胜任的
chromatic aberration	色差
clockwise	顺时针方向的
collimated	准直
coma	彗差
cometlike	彗星状的
concave	凹的
conjugate points	共轭点
convention	习惯，公约，协定
convex	凸的
counterclockwise	逆时针方向的
critical angle	临界角
cross section	横截面
cumbersome	麻烦的，不方便的
defect	缺点，缺陷，瑕疵，损伤
derivations	引出
dispersion	色散
dissimilar	不相似的
distortion	畸变
durable	耐久用的，经久的，坚固的
erect image	正立像
exhibit	呈现，陈列，展出
field curvature	场曲
formalism	体系
geometry	几何学
give rise to	引起，产生，导致
go through	通过
in connection with	与……有关，关于
incident	入射的
inclination	倾斜（角），偏角，倾向
index of refraction	折射率

indices	index 的复数
infinitely	无穷，无限
intact	完整的，原封不动的
intense	强烈的
intercept	截取，拦截，相交，折射
intersect	相交，贯穿
invert image	倒立的像
isosceles	等腰
lest	以免
magnification	放大率
magnitude	数量级
medium	介质
meniscus	弯月形
metallic	金属(制)的
numerically	在数值上
oxidation	氧化
parallel	平行的，类似的
parameter	参数
parenthetical	括号中的
parenthetical	附加的
peculiar	特殊的
polish	抛光，擦亮
primary focal point	主焦点
prism	棱镜
propagate	传播
property	特性，特征
quote	引号
radius	半径
rather	相反地，反而，倒不如说（在句首，或作插入语，其前通常是否定句）
ratio	比，比值
real image	实像
reciprocal	倒数
rectilinear	直线运动的
reflection	反射
refraction	折射
respectively	分别地，各自地
scratch	刮伤，擦伤
secondary focal point	副焦点
sharpness	锐度

single-element lenses	简单透镜
spectrum	谱，光谱
spherical aberration	球差
subscript	下标记
successive	连续的，逐次的，递次的
aperture stop	孔径光阑
paraxial approximation	傍轴近似
tangent	切线
transparent	透明的，半透明的
trigonometry	三角法
universally	一般地，普遍地
vacuo	（拉丁语）真空
vanish	消失，消散
virtual image	虚像
with regard to	关于，论及，对于，就……而言
withstand	抵抗，经得起，经受住

Grammar　专业英语翻译方法（一）：英汉句法对比的总结

英语：

1．主、谓结构严明，动作行为都有主语；
2．多被动语态；
3．介词繁多，名词亦多，应用广泛；
4．复合句多用“形合法”，根据主、宾、定、状等语法和句法的关系与短语组合成句子，繁而不乱；
5．语序灵活多变，纵横交错，重点突出，结构严谨，主次分明；
6．属综合性语言，着重词形、人称、时态、语态、语气的变化；
7．句法结构复杂，多长句。

汉语：

1．主、谓结构往往不全，常见无主语句或无人称句；
2．很少用被动语态；
3．多用动词，少用介词；
4．复合句多用“意合法”，句子成分很少用连词，根据事理演变或发展过程和逻辑关系，靠语意串联，承上启下，一气呵成；
5．语序自然，相映成趣，结构灵活，词句简洁，有如修竹，节节有序；
6．属分析性语言，根本无词形变化，重意义而不重形态；
7．句法自然，注意修辞，多简单句型。