(Appendix) Where do the selection rules originate from and what is the dipole approximation?
The following paragraph offers a simplified explanation of the derivation of selection rules. It should, however, be emphasised that not all simplifying assumptions are necessarily mentioned explicitly.
According to Fermi's golden rule, the intensity of a transition between two states $|i\rangle$ and $|f\rangle$ ('final state' and 'initial state') is generally proportional to the square of a matrix element

$$I \sim |M_{fi}|^2$$

where the matrix elements are those of the electromagnetic interaction operator

$$M_{fi} \sim \langle f \,|\, \mathcal{H}_{WW} \,|\, i \rangle$$

The electromagnetic interaction operator of a non-relativistic 1 electron-1 photon process can be derived from the integral

$$\mathcal{H}_{WW} = -\frac{1}{c} \int \mathrm{d}^3 r \hat{j}(\vec{r}) \hat{A}(\vec{r})$$

Here, $\hat{j}(\vec{r})$ is the current density operator of the electron; $\hat{j}$ leads to a term $\sim \frac{e}{2m_e}\hat{p} \sim \hat{p}$ and is therefore proportional to the momentum operator of the electron.
$\hat{A}(\vec{r})$ is the vector potential eventually leading to a term $\sim \vec{\epsilon} e^{i\vec{k}\vec{r}}$, where $\vec{\epsilon}$ is the polarisation and $\vec{k}$ is the photon wave vector.
Inserting the respective values thus yields the expression

$$M_{fi} \sim \left\langle f \left| e^{i\vec{k}\vec{r}} \hat{\vec{\epsilon}} \hat{\vec{p}} \right| i \right\rangle$$

This matrix element now needs to be evaluated.
The term $e^{i\vec{k}\vec{r}}$ can be simplified using a Taylor series expansion:

$$e^{i\vec{k}\vec{r}} = \underbrace{1}_{\text{yields el. dipole term (E1)}} + \underbrace{i\vec{k}\vec{r} + o(i\vec{k}\vec{r})^2}_{\text{yields magnetic dipole moment M1, electric quadrupole moment E2 etc.}}$$

The electric dipole radiation usually dominates by far (if it is allowed!), as can be shown by a simple estimation of an optical transition ($\lambda \simeq 6000\,\text{Å}, r \simeq 1\,\text{Å}$)

$$kr \simeq \frac{2\pi}{6000\,\text{Å}} 1\,\text{Å} \simeq \underbrace{\frac{1}{1000}}_{\text{M1, E2 etc}} \ll \underbrace{1}_{E1}$$

In this example, the amplitudes (E1 and M1, E2 etc) differ by 3 orders of magnitude ($10^3 : 1$) and this difference is even squared in the case of their ratios ($\simeq 10^6 : 1$).
In dipole approximation, the matrix element to be evaluated is therefore

$$M_{fi} \sim \langle f \,|\, \vec{\epsilon}\vec{p} \,|\, i \rangle$$

The operator $\hat{\vec{\epsilon}}\hat{\vec{p}}$ can be simplified:

$$\begin{aligned}
\langle \Phi_m \,|\, \vec{\epsilon}\vec{p} \,|\, \Phi_n \rangle &= \vec{\epsilon} \left\langle \Phi_m \left| m_e \frac{\mathrm{d}\vec{r}}{\mathrm{d}t} \right| \Phi_n \right\rangle \\
\text{(Time derivative of an operator)} \qquad &\sim m_e \vec{\epsilon} \left\langle \Phi_m \left| \frac{i}{\hbar} [H_0 \,,\, \vec{r}] \right| \Phi_n \right\rangle \\
([H_0 \,,\, r] \sim p, \text{ see Bethe/Salpeter, p.52}) \qquad &\sim \frac{i m_e}{\hbar} \vec{\epsilon} \langle \Phi_m \,|\, H_0 r - r H_0 \,|\, \Phi_n \rangle \\
(\Phi_i \text{ are eigenstates of } H_0) \qquad &\sim i m_e \vec{\epsilon} \left(\frac{E_m - E_n}{\hbar} \right) \langle \Phi_m \,|\, \vec{r} \,|\, \Phi_n \rangle \\
&\sim -i m_e \omega \langle \Phi_m \,|\, \vec{\epsilon}\vec{r} \,|\, \Phi_n \rangle
\end{aligned}$$

Therefore, in the present context

$$\vec{\epsilon}\vec{p} \sim \vec{\epsilon}\vec{r}$$

This corresponds to a dipole energy $\sim \vec{E}\vec{\mu} = q\vec{E}\vec{r} \sim r\cos\vartheta$.
Thus, the selection rules for dipole radiation (E1) can be regarded as conditions under which

$$\langle f \,|\, \vec{\epsilon}\vec{r} \,|\, i \rangle \neq 0$$

Application of the calculation to an H atom
If we perform the calculation for atoms similar to H with the known energy eigenstates $\Phi_m = R_{n,l}(r)Y_{l,m}(\vartheta,\varphi)$, we obtain

$$\langle \Phi_m \,|\, \vec{\epsilon}\vec{r} \,|\, \Phi_n \rangle = \underbrace{\int_0^\infty \mathrm{d}r r^2 R^*_{n_f,l_f}(r) r R_{n_i,l_i}(r)}_{\text{radial part, const.}} \underbrace{\int\int \mathrm{d}\varphi\, \mathrm{d}\vartheta \sin\vartheta Y^*_{l_f,m_f}(\vartheta,\varphi) Y_{l_i,m_i} \vec{\epsilon}\frac{\vec{r}_e}{|r|}}_{\text{angular part, is evaluated in the following}}$$

Expressing $\vec{\epsilon}\vec{r}$ via spherical surface functions $Y_{l,m}$

$$\begin{aligned}\vec{\epsilon}\vec{r} &= \epsilon_x \sin\vartheta\cos\varphi + \epsilon_y \sin\vartheta\sin\varphi + \epsilon_z\cos\vartheta \\ &= \sqrt{\frac{4\pi}{3}}\left(\epsilon_z Y_{1,0} + \frac{-\epsilon_x + i\epsilon_y}{\sqrt{2}} Y_{1,1} + \frac{\epsilon_x + i\epsilon_y}{\sqrt{2}} Y_{1,-1}\right)\end{aligned}$$

shows that the matrix element needs to be evaluated using an integral over three $Y_{l,m}$.

$$0 \overset{!}{\neq} \text{angular part} = \int\int \mathrm{d}\vartheta\, \mathrm{d}\varphi \sin\vartheta Y^*_{l_f,m_f} Y_{(l=1)(m=-1,0,1)} Y_{l_i,m_i}$$

This expression is only $\neq 0$ if:

$$m_f - m_i = \pm 1 \qquad \text{and} \qquad l_f - l_i = 0, \pm 1$$

(The selection rule $\Delta S = 0$ is due to the fact that $\mathcal{H}_{WW}$ does not contain — and therefore cannot change — the spin.)

References

- Bethe/Salpeter
- Gasiorowicz