Мы знаем, что $y = f(x) + \epsilon$, где ϵ - независимая с x случайная величина с нулевым математическим ожиданием и некоторой дисперсией σ^2. Мы строим алгоритм определения y по x $a_X(x)$, обучаясь на данных X. Нам нужно найти $E(a_X(x) - y)^2$. Перейдем от математического ожидания случайной величины к математическому ожиданию условного математического ожидания.

$$E(a_X(x) - y)^2 = EE((a_X(x) - y)^2|x)$$

Воспользуемся свойствами условного математического ожидания, чтобы преобразовать это выражение:

$$E((a_X(x) - y)^2|x) = E((a_X(x) - f(x) - \epsilon)^2|x) =$$
$$= E((a_X(x) - f(x))^2 - 2\epsilon(a_X(x) - f(x)) + \epsilon^2|x) =$$
$$= (a_X(x) - f(x))^2 - 2(a_X(x) - f(x))E(\epsilon|x) + \sigma^2 =$$
$$= E((a_X(x) - f(x))^2|x) + \sigma^2 =$$
$$= E((a_X(x) - Ea_X(x) - (-Ea_X(x) + f(x)))^2|x) + \sigma^2 =$$
$$= E((a_X(x) - Ea_X(x))^2 + (f(x) - Ea_X(x))^2 - 2(a_X(x) - Ea_X(x))(f(x) - Ea_X(x))|x) + \sigma^2 =$$
$$= (a_X(x) - Ea_X(x))^2 + (f(x) - Ea_X(x))^2 + \sigma^2$$

Таким образом, учитывая что $f(x) = E(y|x)$, $\sigma^2 = E(y - E(y|x))^2$, имеем:

$$E(a_X(x) - y)^2 = E(a_X(x) - Ea_X(x))^2 + E(E(y|x) - Ea_X(x))^2 + E(y - E(y|x))^2 =$$

$$= Variance + Bias + Noise$$

Что и требовалось доказать.