Fabrication

The Z80 was made using a NMOS fabrication process. The process requires many steps of processing to form the chip.

The key idea of chip fabrication is to put a protective layer over part of the chip and then expose the chip to a chemical process. This will modify the chip in the unprotected area, but not in the protected area. The protective layer is removed and the process is repeated. This builds up complex layered structures on the chip. The layers are usually separated by insulating silicon dioxide.

To form the protective layer, a light-sensitive chemical called photoresist is put on the chip. Light then shines through a patterned mask onto the photoresist. Where light hits the photoresist, the photoresist hardens. The photoresist is then dissolved off, leaving only the hardened photoresist behind. Thus, the photoresist pattern on the chip matches the pattern of the mask.

By looking at the die of the chip, we can see marks corresponding to each of the masks. The Z80 appears to have used 7 masks, which was a typical number for an NMOS chip like the Z80.

Mask 1: Silicon dioxide mask. This defines the areas of silicon that will be subject to diffusion, even though that doesn't happen until later.

Mask 2: Ion implantation. This mask selects areas that have ion implantation, which forms the depletion pull-up transistors, as well as the “traps”.

Mask 3: Buried contact. This mask defines the buried contacts, contacts between silicon and polysilicon.

Mask 4: Polysilicon. This mask defines the position of polysilicon, including the gates.

Mask 5. Contact cuts: This mask defines the position of contacts between

metal and polysilicon or silicon.

Mask 6. Metal. This mask defines the position of the metal layer.

Mask 7. Overglass. This mask defines the area of passivation over the chip. The pads are left uncovered.

The following will describe a typical NMOS process (ref: VLSI Design techniques for analog and digital circuits). The actual process used for the Z80 may differ in the details but will be similar. Reference: http://www.iaa.ncku.edu.tw/~aeromems/MEMSDesign/Ch2.pdf

The process starts with a polished p-type silicon disk, typically 500 micrometers thick. The wafer is oxidized to add a SiO2 layer of typically 1 micrometers.

Photoresist is added on top of this and exposed to light through Mask 1. The photoresist is removed from areas that are not to be diffused. The chip is then etched to remove the silicon dioxide in the unprotected areas. The remaining (hardened) photoresist is removed.

Photoresist is added and exposed to light through Mask 2 to define the regions for depletion transistors. The unexposed photoresist is removed. Arsenic atoms are used to dope the depletion regions using ion implantation. The remaining photoresist is removed.

A thin layer of silicon dioxide of .1 micrometer is grown over the entire surface. This silicon will form the gate insulation for transistors.

Photoresist is added and exposed to light through Mask 3 to define the buried contacts. The unexposed photoresist is removed. The chip is etched to remove the silicon dioxide from the unprotected regions. This leaves the silicon exposed where buried contacts will be formed. The remaining photoresist is removed.

A polysilicon layer is grown over the entire surface. This layer is created by

chemical vapor deposition (CVD). It will be separated from the silicon by the thin silicon dioxide region, except where it directly contacts the silicon to create a buried contact.

Photoresist is added and exposed to light through Mask 4 to define the polysilicon. The unexposed photoresist is removed. The chip is etched to remove the polysilicon from the unprotected regions, leaving the polysilicon wires and gates. The remaining photoresist is removed.

Photoresist is added and exposed to light through Mask 4 to define the polysilicon. The unexposed photoresist is removed. The chip is etched to remove the polysilicon from the unprotected regions, leaving the polysilicon wires and gates. The unprotected thin oxide is also removed by etching. The remaining photoresist is removed.

At this point, the chip is protected by thick oxide except where there are openings as defined by Mask 1. Polysilicon wires run over the chip. Where gates will be formed, the polysilicon is separated from the silicon by thin silicon dioxide.

The wafer is heated to a high temperature and n-type impurities are diffused into the chip by exposing it to a gas containing, for example, phosphorus. This is the step where transistors are formed. The gates are self-aligned with the source and drain because the polysilicon protects the gates from diffusion. The self-aligning nature of the transistors is a key innovation in NMOS technology.

A thick silicon dioxide layer is grown over the entire chip. Photoresist is added and exposed to light through Mask 5 to define the contact cuts, where the metal contacts the silicon or polysilicon. In the unprotected regions, the silicon dioxide is etched away. The remaining photoresist is removed.

The chip has 1 micrometer of aluminum deposited on the surface to form the metal layer. Photoresist is added and exposed to light through Mask 6 to define the metal layer. In the unprotected regions, the metal is etched away. The remaining photoresist is removed.

The chip has a thick layer of silicon dioxide overglass added to protect the chip surface. Photoresist is added and exposed to light through Mask 7 to define the pads. In the unprotected regions, the silicon dioxide is etched away. The remaining photoresist is removed.

This completes the manufacture of the chip wafer. The dies are tested while still part of the wafer, cut into dice and packaged to make the final integrated circuits.