

# 集成电路中的物理问题讲座

## 第十讲 集成电路技术的一些物理限制

马俊如

(中国科学院半导体研究所)

1958年第一个平面型晶体管的诞生,为集成电路的发明奠定了坚实的技术基础。翌年,集成电路问世,其后集成规模以每年翻一番的高速度迅速发展。现已有能在几毫米见方的小硅片(又称芯片)上集成包含百万只元器件的超大规模集成电路。这二十多年中,一个芯片中集成的元器件数增加了几十万倍,而芯片的面积仅扩大了一个数量级左右。因此集成技术的发展主要是依靠缩小小芯片中每个器件的尺寸来实现。实际上器件的线度缩小了近三个数量级,并已采用1—3微米宽的线条尺寸作为器件制造的设计规则。人们指望进一步缩小器件的尺寸以获得超超大规模集成电路,这种努力正在进行中。

集成电路缩小器件尺寸所展现的巨大优越性,已在磁泡、超导电子器件和集成光学、表面声波器件等方面得到了成功的应用。这就吸引人们花费大量的人力、物力来研究如何大幅度缩小器件的尺寸。另一方面,缩小尺寸的技术难度非常大,到一定阶段会遇到两方面带根本性的限制,即器件运用物理原理方面的限制和人们掌握的工具手段和技术的能力限制。本文将着重讨论发展超大规模集成电路的主要器件——MOS晶体管及工艺方面的一些限制。

### 一、按比例缩小原理<sup>[1]</sup>

集成电路中的MOS晶体管结构见图1。它是在硅衬底上制作具有一定形状的金属层、多晶硅层、绝缘SiO<sub>2</sub>层和高掺杂层(由扩散或离子注入形成的导电层)组成。栅区没有电荷

时,源与漏间是处于反向的两个背靠背p-n结,没有电流,不导通。若栅区上有足够多的正电荷存在,栅与源间的外加电压V<sub>gs</sub>大于使源与漏间有电流的最小电压V<sub>th</sub>,则是导通状态。MOS晶体管的基本原理就是用栅上的电荷来控制源漏间的电流,从而实现开关功能。

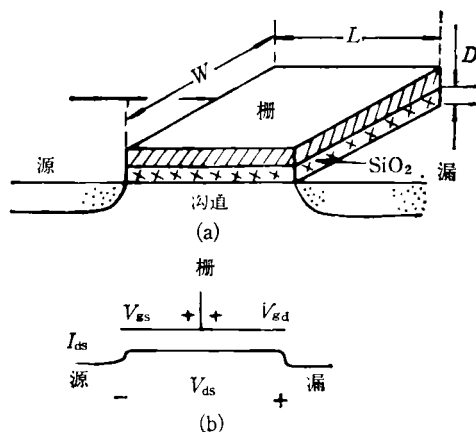


图1 MOS晶体管结构和符号示意

(a) MOS晶体管结构; (b) MOS晶体管的符号表示

设在源与漏间加一个电压V<sub>ds</sub>,则电子从源区到达漏区所需的过渡时间τ为

$$\tau = \frac{L^2}{\mu V_{ds}}, \quad (1)$$

其中μ是电子的迁移率。栅是由厚度为D的SiO<sub>2</sub>绝缘层把它与硅衬底隔开,所以形成了一个电容C<sub>g</sub>。

$$C_g = \frac{\epsilon W L}{D}. \quad (2)$$

沟道中流过的电流可以表示为

$$I_{ds} = \frac{\mu \epsilon W}{L D} (V_{gs} - V_{th}) \cdot V_{ds}. \quad (3)$$

如果把器件的线度尺寸缩小  $n$  倍 ( $n > 1$ ), 即新的器件尺寸表示为

$$\begin{aligned} D_n &= D/n, \\ L_n &= L/n, \\ W_n &= W/n. \end{aligned} \quad (4)$$

在器件运用中所选的电场接近击穿或隧道漏电出现时的最大值, 因此在缩小尺寸时电场保持不变, 即缩小电压  $V_n$ ,

$$V_n = V/n. \quad (5)$$

由上列各式可推导出各参量的相应变化:

$$\tau_n = \tau/n, \quad (6)$$

$$C_n = C/n, \quad (7)$$

$$I_n = I/n. \quad (8)$$

器件的开关功率  $P_{sw}$ , 直流功率  $P_{dc}$  和开关能量  $E_{sw}$ , 分别推导如下:

$$P_{sw(n)} = \frac{C_n V_n^2}{2\tau_n} = \frac{P_{sw}}{n^2}, \quad (9)$$

$$P_{dc(n)} = I_n V_n = \frac{P_{dc}}{n^2}, \quad (10)$$

$$E_{sw(n)} = \frac{1}{2} C_n V_n^2 = \frac{E_{sw}}{n^2}. \quad (11)$$

从(5)–(11)式可以清楚看出, 缩小器件尺寸对性能有利。例如工作速度加快了, 单位面积中可集成的器件数目增加了, 功率密度保持不变, 不利的是电流密度增加了。按比例缩小的原理推论下去, 原则上 MOS 晶体管可以尽量缩小。但从器件结构和性能的进一步分析表明, 存在一系列限制, 不可能无限制地缩小, 下面讨论最基础的限制。

## 二、MOS 晶体管的最小尺寸

控制 MOS 晶体管结构的尺寸, 最关键的是沟道长度  $L$  和栅氧化层的厚度  $D$ 。因此缩小 MOS 管的尺寸, 主要就是缩小  $L$  和  $D$ 。

从器件的结构考虑, 缩小  $L$  的主要限制是源与漏 p-n 间的穿透, 即两个结的耗尽层相碰。由图 1, 当  $L=0$  时, 源与漏就直接连通了。对于突变结, 耗尽层厚度  $T$  可以写为

$$T = \sqrt{\frac{2\epsilon(V + \phi)}{qN_B}}, \quad (12)$$

其中  $V$  是外加电压,  $\phi$  是结自建电势差,  $N_B$  是 p-n 结低掺杂一边的掺杂浓度。对于给定的  $N_B$ ,  $T \propto V^{1/2}$ , 即  $T$  随外加反偏压而增加。  $T$  的增加主要在低掺杂一边。对实际的 MOS 结构是出现在沟道区, 使有效沟长缩短。由此得知, 最小沟长  $L_{min}$  是指最大反偏压下源、漏结耗尽层不相碰, 即

$$L_{min} = 2T_{max}, \quad (13)$$

式中 2 倍是指源和漏两个结耗尽层厚度之和, 这里假设两个结相同。从 (12) 式又得知  $T \propto N_B^{-1/2}$ 。其物理意义是, 增加  $N_B$ ,  $T$  就减小, 这是由于单位耗尽层宽度内的空间电荷增加了, 较小的  $T$  足以形成结中平衡的自建场。所以随着  $N_B$  增加,  $L_{min}$  减小,  $N_B$  也不能随意增加。在 MOS 结构中, 限制  $N_B$  最高值有两个因素:

1. 场发射: 正常 p-n 结的反向只有少数载流子形成的极微弱的饱和电流, 与正向导电电流相比, 可以认为是截止状态。当  $N_B \geq 2 \times 10^9 \text{cm}^{-3}$ , 量子隧道的场发射显著, 反向和正向特性类似, 源与漏间的绝缘消失。它限制了  $N_B < 2 \times 10^{10} \text{cm}^{-3}$ 。

2. 氧化层的击穿: 隔开栅电极与沟道是一层厚度为  $D$  的  $\text{SiO}_2$  层。  $\text{SiO}_2$  的击穿场强近似为  $6 \times 10^6 \text{V/cm}$ 。这个场强不能使掺杂浓度  $N_B > 1.3 \times 10^{19} \text{cm}^{-3}$  的沟道反型, 变为可以使电子流过的通道。所以  $\text{SiO}_2$  的击穿限制  $N_B \leq 1.3 \times 10^{19} \text{cm}^{-3}$ 。

由这两点可确定

$$(N_B)_{max} \leq 1.3 \times 10^{19} \text{cm}^{-3}.$$

在这个浓度下, 当无外加偏压时,  $T = 0.01 \mu\text{m}$ 。即从物理上讲, 必须使  $L_{min} \geq 0.02 \mu\text{m}$ , 否则源与漏结不加反偏压就穿通了。如果联系具体结构设计, 分析表明,  $L_{min} \approx 0.2 \mu\text{m}$ 。目前已经达到的水平是  $1 \mu\text{m}$ 。

关于栅下  $\text{SiO}_2$  层  $D$  的限制, 从原理上分析主要由量子效应决定。夹于两导体中间非常薄的绝缘膜, 若在其一边加上电子波函数, 到另一

边并不完全衰减为零。电子总有一定的几率由量子隧道穿过介质膜,即有一定的电流<sup>[2]</sup>。分析表明, SiO<sub>2</sub> 薄膜导电的机构主要是接触而不是体内<sup>[3]</sup>。在 SiO<sub>2</sub> 与单晶、多晶硅或金属的接触界面处,电子有一定的几率穿过 SiO<sub>2</sub> 的禁带进入到 SiO<sub>2</sub> 的导带中。SiO<sub>2</sub> 单位面积上的电导随膜厚的减薄呈指数式上升,即绝缘变坏,由此得知, SiO<sub>2</sub> 层的厚度  $D$  所能达到的最小值是指通过 SiO<sub>2</sub> 薄层的隧道电流,比维持 MOS 管正常工作的电流小到可以忽略不计时所定义的厚度,约为 5nm。目前实际上已用 10—20nm 的 SiO<sub>2</sub> 层,它与极限相差不远了。

### 三、最小尺寸工艺原理的限制

一个器件从原理上所允许达到的最小尺寸和在加工中能允许做到的最小尺寸同样是令人关心的问题。原则上讲,集成电路的结构是在固体上加工,因此可加工的最小尺寸是构成固体材料的原子线度<sup>[4]</sup>,即几个埃。但各种加工方法在原理上都有各自的限制,往往其极限比原子的线度大得多。下面从工艺原理的角度,对小尺寸的限制作一些讨论。

#### 1. 线条边缘的误差

可用于微小尺寸图形成形的工艺主要是紫外线、X 射线、电子束和离子束加工技术等。这里我们考虑加工一物理线度  $L$ ,它可以表示光刻胶的线度宽度、硅衬底上 SiO<sub>2</sub> 窗口宽度或金属互连条宽等。从量子力学观点看,  $L$  边缘的误差  $\Delta L$  与加工所用束中粒子的动量变化  $\Delta p$  符合海森堡测不准关系式:

$$\Delta p \cdot \Delta L \geq h. \quad (14)$$

设动量的最大变化为  $+p \rightarrow -p$  或  $\Delta p = 2p$ <sup>[5]</sup>, 则有

$$\Delta L = \frac{h}{2p}, \quad (15)$$

这里  $\Delta L$  的物理意义是,在线条  $L$  的边缘不可能做到比  $\Delta L$  更好。如果考虑的线条就是  $\Delta L$ , 则它是最小的尺寸。

实际上,要定义一物理线条  $L$ ,其误差必须

考虑它的两端。若用束来操作加工,第一次操作是确定一端的误差  $\Delta L$ ,接着第二次操作是另一端的误差  $\Delta L$ 。设这二次操作是独立无关,都以一个参考点为依据(例如对准标记等)由于这个参考点本身也同样有误差  $\Delta L$ ,它直接影响了  $L$  两端的误差。 $L$  每一端相对于参考点的误差为  $\sqrt{2} \Delta L$ ,则两端相对的误差为  $2\Delta L$ 。这是一维考虑的结果,对于二维的情况是  $2\sqrt{2} \Delta L$ ,三维时为  $2\sqrt{3} \Delta L$ <sup>[6]</sup>。作为基本物理限制的讨论,我们仅考虑一维的情况。下面以光学和电子束加工为例对最小尺寸作一些描述。

#### 2. 光学加工的最小尺寸

光学加工使用紫外到软 X 射线波段,波长为 450—0.4nm。实际应用中,考虑多方面因素,并不是全波段内进行,而是分波段,即 450—260nm 和 4—0.4nm。对于光子,

$$p = \frac{h}{\lambda}, \quad (16)$$

代入(15)式,则有

$$\Delta L \geq \frac{\lambda}{2}. \quad (17)$$

这就是通常的光衍射极限,波长  $\lambda$  愈短,  $\Delta L$  愈小。紫外光  $\lambda \approx 400\text{nm}$ ,  $\Delta L \geq 0.2\mu\text{m}$ ; 软 X 射线  $\lambda = 1\text{nm}$  时,  $\Delta L \geq 0.5\text{nm}$ 。这正是为什么人们认为加工到线条宽度小于  $1\mu\text{m}$  时,紫外光刻快达到极限,而在光的应用上需要开发软 X 射线的物理背景。

#### 3. 电子束曝光的工艺原理限制

粒子具有波动性,其德布罗意波长为

$$\lambda = \frac{h}{\sqrt{2mE}}. \quad (18)$$

由(15)式得到  $\Delta L$  为

$$\Delta L = \frac{h}{2} (2mE)^{-1/2}. \quad (19)$$

对于电子,  $m = 9.1 \times 10^{-31}\text{kg}$ , 设能量  $E$  为 1500eV, 则波长为 0.01nm,  $\Delta L \geq 0.005\text{nm}$ 。而且,能量越高,  $\lambda$  越短,  $\Delta L$  越小。但这仅仅是理想的结果,实际上我们必须要考虑电子束在与光刻胶、硅等靶材料相互作用时的散射以及

胶的颗粒度对  $\Delta L$  的限制。

### (1) 电子束在靶中的散射

束扩张的主要原因是束中电子在靶中与原子相互作用引起的散射。电子的散射分为两种情况，沿电子运动方向偏离小于  $90^\circ$  的前向散射和与运动方向相逆的背散射。电子大部分是前向散射<sup>[7]</sup>。设有一均匀的能量为  $E$  的平行电子束，打到靶上，在能量不太高的情况下，可以认为散射是随机的，即每次碰撞与前次无关。因散射而引起的束扩张偏角  $\Delta\phi$  为<sup>[8]</sup>

$$\Delta\phi = \frac{4 \times 10^5}{E} \cdot Z \cdot \left(\frac{\rho x}{A}\right)^{1/2}, \quad (20)$$

其中  $\rho$  是靶材的密度， $Z$  是原子序数， $x$  是穿透深度， $A$  是原子量。设  $E = 1000\text{eV}$ ， $x = 17 \text{ \AA}$ ，对于硅靶， $\Delta\phi = 36^\circ$ 。则散射的圆锥张角为  $72^\circ$ 。

### (2) 光刻胶的颗粒度

在集成电路加工中，人们最感兴趣的是电子束对光刻胶的轰击，即曝光。对曝光有贡献的是电子的非弹性散射。电子将能量传递给胶分子，使它发生化学结构变化，物理性能与原先不同，例如溶解速率相差很大，从而显现反差。

为得到小的  $\Delta L$ ，必须降低散射引起的束扩张。很显然，需要减薄光刻胶的厚度。光刻胶的分子能够使曝光结构特性产生变化的最小直径为  $0.7\text{--}1 \text{ nm}$ 。因此最薄是一个单层，即  $0.7\text{--}1 \text{ nm}$  厚，但这样的胶容易有漏孔，完整性不好，若是三个单层交叠，则为  $2.5 \text{ nm}$  左右。由 (20) 式推断出的  $\Delta L \approx 2.0\text{--}2.5 \text{ nm}$ 。比测不准关系引起的  $\Delta L$  大两个数量级。

### (3) 邻近效应

通常见到的电子束曝光的光刻胶图形如图 2 所示。其外边角缺少，而相邻两线条间又扩张出去，统称为邻近效应。这也是电子散射引起的严重后果。图中的  $A$  点接受的曝光剂量比边界上的  $B$  点大一倍，因为  $A$  点可以接受四周来的散射电子。 $C$  点的剂量相当于  $A$  点的  $1/4$ 。因此在同样曝光条件下，边角曝光显得不足。对于线条  $l$ ，则两端因曝光不足而缩进去，两边则变细。从探讨最小尺寸的限制而言，这

种情况是无害的。另一方面，从图中也可看到，在两线条非常靠近时，其相邻处的部分突出了。其原因可用图 3 解释，它描述了电子在靶中沿  $x$  方向的几率分布。如果仅考虑第一条线，并当标准偏差为  $\sigma$  时，设其为正态分布，边界在  $x_1$ ；当有第二条线存在时， $x_1$  以内的线条区变化不大，而间隔区中的概率增加了一倍，因而线条变宽了，即造成突出的图形。分析表明，考虑邻近效应，最小尺寸不能小于  $10 \text{ nm}$ 。

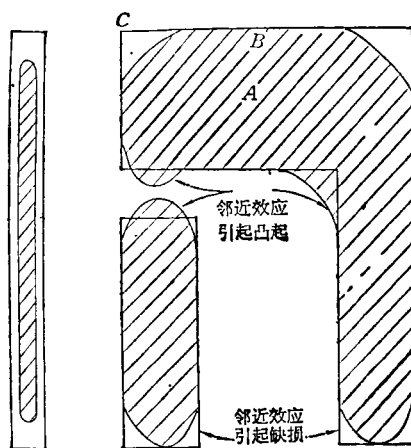


图 2 邻近效应的两种情况

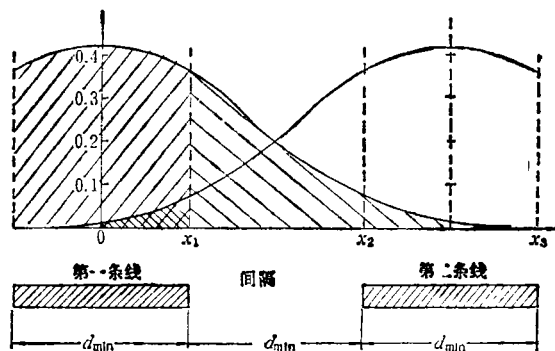


图 3 电子束打靶的概率，设其为正态分布  
(两线条和间距相等为  $d$ )

综上所述，用电子束曝光产生图形，对单个线条的线宽可达到  $2 \text{ nm}$ ；当有线条组合时，线宽的最小值约为  $10 \text{ nm}$  左右。离子束曝光的最小线宽，其量子力学理论限制原则上与电子束相

似,只是离子质量比电子大得多,因而  $\Delta L$  更小.图4列出了各种束的最小线宽  $\Delta L$  与束粒子能量的关系.

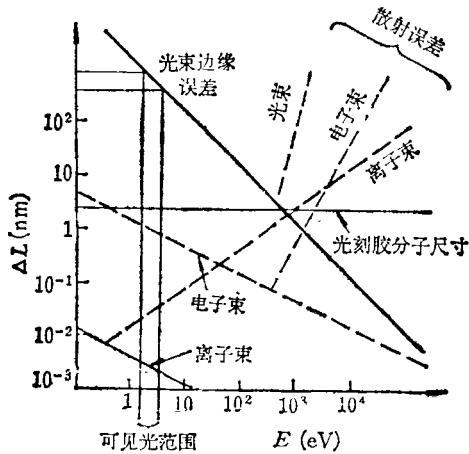


图4 光束、电子束、离子束的  $\Delta L$  与能量  $E$  的关系

#### 四、材料的一些限制

集成技术中使用的半导体、绝缘物和金属材料特性,直接影响器件的性能和应用水平.下面对过渡时间和电迁移作一些简略的讨论.

##### 1. 材料对过渡时间的限制

介质的击穿现象,限制了在半导体材料中加的最大电场,从而限制了器件的运用速度.硅中临界场强  $E_c = 3 \times 10^5 \text{V/cm}$ , 电子的最大速度  $v_{\max} \cong 8 \times 10^6 \text{cm/s}$ , 则在  $\Delta z$  线度内的过渡时间  $\Delta t$  为

$$\Delta z = v_{\max} \cdot \Delta t. \quad (21)$$

由于在  $\Delta z$  两端的电势  $\Delta V$  与  $E_{\max}$  的关系为

$$\Delta V = E_{\max} \cdot \Delta z, \quad (22)$$

则由(21)和(22)式可得

$$\Delta t = \frac{\Delta V}{E_{\max} \cdot v_{\max}}. \quad (23)$$

由于信息的每个基本动作的最小热力学的能量为  $kT^{[9]}$ , 最小的  $\Delta V_{\min} = \frac{kT}{q}$ , 则由(21)式得

到最小的过渡时间  $\Delta t_{\min} = 10^{-14} \text{s}$ .

##### 2. 电迁移效应

电迁移现象是一种扩散过程,它的产生是

由于固体中的原子在外力的干扰或化学势梯度的干扰下,使其从一个地方移动到另一个地方.当一导体中有原子流通过而且散开时,则会引起导体的形状变化并导致烧毁.电迁移对集成电路的影响,通常是指在强电流下金属互连线的烧毁断开.要使一连线断开,需移走一定量的材料,其机理有二种:

(1) 热扩散: 各种材料都有点阵空位.当有温度梯度存在时,空位可由一个位置扩散到另一位置.设原子的扩散系数为  $D_a$ , 导体的电阻率是  $\rho$ , 流过的电流为  $i$ , 则在移动的距离为  $d$  时,电流  $i$  与  $d$  有下列关系<sup>[10]</sup>:

$$d^4 > D_a t_L (Q/kT) (\rho i^2 / \lambda T), \quad (24)$$

其中  $Q$  是能量,  $\lambda$  是热导,  $T$  是温度,  $t_L$  是原子流通过  $d^2$  截面的时间. 设  $D_a \approx 10^{-12} \text{cm}^2 \cdot \text{s}$ , 得到的  $d$  与  $i$  关系如图5所示.

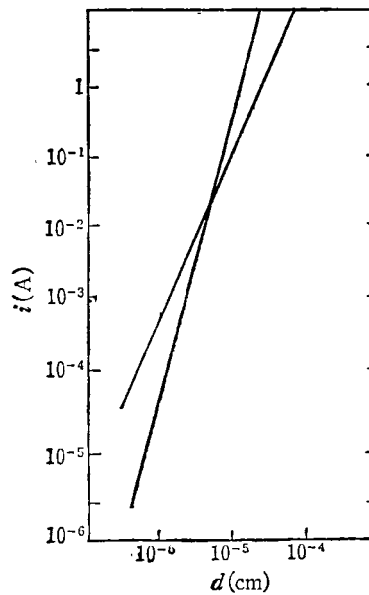


图5 由电迁移现象限制的导体中最大电流与尺寸  $d$  的关系(曲线计算时设  $D_a = 10^{-12} \text{cm}^2/\text{s}$ ,  $q\mu_e = 0.1 \text{cm}^2/\text{N} \cdot \text{s}$ ,  $t_L = 10^8 \text{s}$ ,  $\rho = 10^{-2} \Omega \cdot \text{cm}$ ,  $k = 1.5 \text{W/cm} \cdot ^\circ\text{C}$ ,  $Q = 0.1 \text{eV}$ )

(2) 电学效应: 电迁移中原子的输运过程也可由电效应直接产生.在一导体中,电场  $E$  作用于电子的力  $F$  为

$$F = nqE = \frac{i}{\mu_e d^2}, \quad (25)$$

其中  $\mu_e$  为电子的迁移率。  $F$  作用于晶格使一个原子所受的力为  $F/N$ , 产生的原子流为

$$J = u_a F, \quad (26)$$

其中  $\mu_a$  为原子的迁移率。由此可以推得

$$Nd^3 \geq \left(\frac{\mu_a}{\mu_e}\right) (it_L). \quad (27)$$

$d$  与  $i$  的关系如图 5 所示。实验表明, 在温度为  $175^\circ\text{C}$  和电流密度为  $10^{10}\text{A}\cdot\text{cm}^{-2}$  时, 电迁移就在晶粒边界等位置产生并造成失效。通常限制电流密度  $\leq 10^9\text{A}/\text{cm}^2$ 。

## 五、杂质分布的统计涨落

控制导电类型并不困难, 精确控制杂质分布及其轮廓形状却不容易, 这受到两方面的限制。一是每种掺杂方法本身的限制; 二是微小区域杂质分布的随机涨落。

### 1. 掺杂方法本身的限制

通常用的掺杂方法是热扩散和离子注入。这两种方法所确定的掺杂分布是一种统计分布, 其精确性有一定的限制。实际得到的扩散浓度误差为  $\pm(5-10)\%$ , 而离子注入则为  $\pm 1\%$  左右。这就是说掺杂浓度为  $10^{18}\text{cm}^{-3}$  时, 其涨落为  $10^{16}-10^{17}\text{cm}^{-3}$ 。

### 2. 微区中杂质的统计涨落

器件尺寸降到微米以下, 宏观上认为是均匀的掺杂这一概念, 就不再适用。例如材料的掺杂浓度为  $4 \times 10^{16}\text{cm}^{-3}$  时, 无外加偏置的 p-n 结耗尽层宽度约为  $0.2\mu\text{m}$ 。若以此宽度为边界围一立方体, 其体积为  $(0.2\mu\text{m})^3$ , 其中仅有约 300 个杂质原子。在这种情况下杂质原子的分布不能再看作是均匀的。如果进一步考虑更小的尺寸, 则其中所含杂质原子数更少, 因而随机的涨落更显著。掺杂统计涨落的存在, 直接影响到在这个尺度下的其它性能。象击穿电压就会随位置而变化。原因是击穿电压决定于结中的最高电场  $E_{\text{max}}$ , 其关系为

$$E_{\text{max}} = [2Nq(V + \phi_B)/\epsilon]^{1/2}, \quad (28)$$

其中  $N$  为杂质浓度, 如果  $N$  有涨落, 则结中各处

的最高电场不同, 因而击穿电压也不同。

从按比例缩小的原理, 抽象地推论, MOS 器件可以无限制缩小其几何尺寸。因而, 在一块硅片的单位面积中, 可集成的器件随尺寸的缩小而呈平方增加。器件物理分析表明, MOS 结构的缩小受最小沟道长度  $L_{\text{min}}$  和栅区  $\text{SiO}_2$  厚度  $d$  的限制, 特别是  $L_{\text{min}}$  直接限制了器件的面积。理论上  $L_{\text{min}} = 0.02\mu\text{m}$ , 考虑掺杂和击穿等因素,  $L_{\text{min}} \approx 0.2\mu\text{m}$ 。实际上要达到  $0.2\mu\text{m}$ , 也需要研究解决热电子效应等一系列短沟道物理问题。

在加工能力的分析上, 每一种具体的加工方法所限制的机理有所不同。  $\lambda = 400\text{nm}$  的紫外线, 由衍射决定的最小限  $\Delta L = 0.2\mu\text{m}$ , 电子束散射的限制为  $2.0-2.5\text{nm}$ , 考虑到组合图形时的邻近效应, 则限制为  $10\text{nm}$  左右。

材料的临界击穿限制了  $\text{SiO}_2$  的厚度需  $\geq 5\text{nm}$ , 硅中的临界电场强度限制了在其中传送的电信号最短过渡时间约为  $10^{-14}\text{s}$ 。电迁移限制的最大电流密度  $\leq 10^9\text{A}/\text{cm}^2$ 。尺寸缩小到毫微米量级时必须注意杂质统计涨落所引起的特性偏差, 甚至可以设想若掺杂浓度不高 (例如  $10^{15}\text{cm}^{-3}$ ), 则在  $10\text{nm}$  围成的某些立方体区域中找不到一个杂质。所以杂质在微区中分布的统计涨落有可能成为器件缩小尺寸的一大障碍。

本文的讨论侧重于从基础的角度分析问题, 许多考虑较理想化, 忽略了有联系的其它很多因素, 以便于问题的简化, 实际情况远为复杂。再则器件和工艺的物理限制是较为广泛的课题, 因限于篇幅本文仅讨论了与提高集成度有关的一部分。

作者对黄碧莲同志就一些问题做的有益讨论表示感谢。

## 参 考 文 献

- [1] I. Brodie and J. J. Muray, *The Physics of Microfabrication*, Plenum Press, New York and London, 443—463.
- [2] B. Hoeneisen and C. A. Mead, *Solid State Electronics* 15-7 (1972), 819—829.

(下转第 755 页)