Method and system for refreshing a memory device during reading thereof

A refresh circuit (130) for refreshing a memory device (100) is proposed. The refresh circuit includes: reading means (110-125, 505m, 505r, 515-520) for reading a set of memory cells (MC), the reading means including means (110-125) for applying a biasing voltage having a substantially monotone time pattern to the memory cells and to a set of reference cells (RC) each one having a predefined reference threshold voltage, means (505m, 505r) for reading the stored state of a certain memory cell by detecting the reaching of a comparison current (Ic) by a cell current (Im) of each memory cell and by a reference current (Ir) of each reference cell, and means (515-520) for determining a refresh condition of said memory cell by again reading the stored state at a later time, and writing means (555) for applying a writing voltage to said memory cell; the refresh circuit further includes control means (505g, 525-550; 525', 705, 530-550) for enabling the writing means during at least part of the application of the biasing voltage after the determination that said cell is in the refresh condition.
Description

[0001] The present invention relates to the memory field. More specifically, the present invention relates to the refresh of memory devices.

[0002] Memory devices are commonly used to store data in a number of applications; particularly, in a non-volatile memory device the data is preserved even when a power supply is off. Typically, the memory device includes a matrix of memory cells (for example, consisting of floating gate MOS transistors). The memory cells are programmed by injecting electrical charge into their floating gates; conversely, the memory cells are erased by discharging the floating gates. The electrical charge in the floating gate of each memory cell defines different levels of its threshold voltage representing corresponding logical values. Particularly, in a multi-level memory device each memory cell can take more than two levels (and then store a plurality of bits).

[0003] The logical values stored in memory cells selected by a corresponding address are read by comparing a current flowing through each selected memory cell with the currents provided by reference cells having pre-defined threshold voltages. For this purpose, a suitable biasing voltage is applied to the selected memory cells and to the reference cells. In a specific reading technique (as described in EP-A-1467377, the entire disclosure of which is herein incorporated by reference) the biasing voltage has a ramp-like time pattern. In this case, the selected memory cells and the reference cells turn on at different times (as soon as the biasing ramp reaches their threshold voltages). The temporal order of the turning on of each selected memory cell with respect to the ones of the reference cells uniquely identifies the logical value stored in the selected memory cell. In this way, the precision of the read operation is strongly improved and made independent of most external factors.

[0004] A very critical problem of the memory devices known in the art is their data retention. In fact, the memory cells are subjected to different electrical stresses during operation of the memory devices; these stresses may cause a partial loss of the electrical charge stored in their floating gates, with an undesired erasure of the memory cells that may even cause a change of the logical values stored therein. This problem is particular acute in modern memory devices, wherein the memory cells have very thin oxide layers. Moreover, the problem is exacerbated in multilevel memory devices, wherein the gap between adjacent levels of the threshold voltages is strongly reduced (down to 0.3V); in this case, even a small drift of the threshold voltages of the memory cells (for example, of 0.1V) may change their logical values.

[0005] The above-mentioned problem may be addressed by the use of Error Correction Code (ECC) techniques. The ECC adds redundant control information to blocks of logical values (for example, consisting of pages formed by 64 bits). This control information is used for detecting and correcting (if possible) errors in the pages. This allows tolerating some drifts of the threshold voltages of the memory cells without impairing operation of the memory device.

[0006] Nevertheless, the structures used to implement the ECC involve a considerable waste of space for storing the control information. Furthermore, the operations required to manipulate the control information slow down operation of the memory device. In addition, the data can be written to and read from the memory device at the level of pages only (because of the need to process the corresponding control information).

[0007] A different solution for preventing the loss of data stored in the memory cells is of refreshing their content. For example, it is possible to copy a sector of the memory device onto a volatile memory device, erasing the sector, and then restoring the data from the volatile memory device; however, any power supply cut during the refreshing operation would cause the complete loss of the data stored in the volatile memory device, and then the impossibility of restoring it onto the non-volatile memory device.

[0008] The refresh operation may also be performed individually for each memory cell (directly on the memory device). For this purpose, the threshold voltage of the memory cell is compared with both read references (discriminating the different logical values that can be stored therein) and corresponding guard references (each one slightly higher than a corresponding read reference). In this way, a warning condition of the memory cell can be detected when its threshold voltage is comprised between any read reference and the corresponding guard reference (meaning that the memory cell has lost electrical charge and then it is dangerously approaching the read reference of its logical value). In this case, a soft-program pulse is applied to the memory cell, so as to inject a small amount of electrical charge into its floating gate (until the correct threshold voltage of the memory cell over the guard reference has been restored).

[0009] The refresh operation is typically performed periodically (for example, after a predetermined number of operations executed on the memory device). Alternatively, it has also been proposed to verify the condition of each memory cell whenever it is read; in this case, the memory cell is refreshed immediately after the read operation (when it is necessary). For example, this technique is described in EP-A-1271552.

[0010] In any case, the refresh operation takes a relatively long time. Particularly, the choice of performing it periodically requires stopping operation of the memory device for the whole period that is necessary to refresh all the memory cells; therefore, this procedure may be untenable in most practical situations. Conversely, when the memory cells are refreshed after their reading the corresponding access time is significantly increased (since the memory cells are not available until completion of the refresh operation); this has a detrimental impact on the performance of the whole memory device.

[0011] All of the above limits the application of the re-
fresh operation; therefore, this may impair the reliability of the memory device.

[0012] In its general terms, the present invention is based on the idea of refreshing the memory device during its reading.

[0013] Particularly, the present invention provides a solution as set out in the independent claims. Advantageous embodiments of the invention are described in the dependent claims.

[0014] More specifically, an aspect of the invention provides a refresh circuit (for refreshing a memory device). The refresh circuit includes reading means that is used to read a set of memory cells. The reading means includes means for applying a biasing voltage having a substantially monotone time pattern (such as a ramp) to the memory cells and to a set of reference cells (each one having a predefined reference threshold voltage). Means is provided for detecting the reaching of a comparison current by a cell current (of each memory cell) and by a reference current (of each reference cell). The reading means further includes means for determining a condition of each memory cell, according to a temporal relation of the reaching of the comparison current by the corresponding cell current and the reference currents. Writing means is provided for applying a writing voltage (such as a soft-program pulse) to at least one of the memory cells. The refresh circuit further includes control means for enabling the writing means during the application of the biasing voltage (or a part thereof), after the determination of the condition of each selected memory cell.

[0015] Typically, the refresh operation is only performed on the memory cells that are in a warning condition.

[0016] For this purpose, it is possible to exploit additional reference cells with predefined (guard) threshold voltages.

[0017] A specific implementation of a circuit that can be used to detect the warning condition is also proposed.

[0018] Alternatively, the same result is achieved by delaying the detection of the reaching of the comparison current by each reference current.

[0019] In this case as well, a specific implementation of a circuit that can be used to detect the warning condition is proposed.

[0020] In an embodiment of the invention, each selected memory cell is soft-programmed immediately (after the determination of its condition) up to the completion of the read operation.

[0021] In a different embodiment of the invention, the selected memory cell is still soft-programmed immediately, but for a predetermined period.

[0022] In a further embodiment of the invention, the selected memory cells are always soft-programmed at the same time (towards the end of the read operation).

[0023] Another aspect of the invention proposes a memory device including the refresh circuit.

[0024] A further aspect of the invention proposes a corresponding method.

[0025] The invention itself, as well as further features and the advantages thereof, will be best understood with reference to the following detailed description, given purely by way of a non-restrictive indication, to be read in conjunction with the accompanying drawings, in which:

Figure 1 is a conceptual representation of a memory device in which the solution according to an embodiment of the invention is applicable;

Figures 2 shows an illustrative time diagram explaining a read operation executed on the memory device;

Figure 3a illustrates an exemplary distribution of the threshold voltage of a generic memory cell;

Figures 3b shows an illustrative time diagram explaining a read operation according to an embodiment of the invention;

Figures 4a-4c are illustrative time diagrams of a refresh operation according to different embodiments of the invention;

Figure 5 is a schematic block diagram of a control unit of the memory device that can be used to implement the solution according to an embodiment of the invention;

Figure 6 illustrates a circuit scheme of a pulser of the control unit according to an embodiment of the invention;

Figure 7 is a schematic block diagram of the same control unit according to a different embodiment of the invention; and

Figure 8 illustrates a circuit scheme of another pulser of the control unit according to a different embodiment of the invention.

[0026] With reference in particular to Figure 1, a non-volatile memory device 100 (for example, consisting of a flash EPROM) is illustrated. The memory device 100 is integrated in a chip of semiconductor material, and includes one or more sectors (only one shown in the figure), each one formed by a matrix 105m of memory cells MC. Each memory cell MC consists of a floating gate MOS transistor with a programmable threshold voltage. The memory cell MC in a non-programmed (or erased) condition exhibits a low threshold voltage. The memory cell MC is programmed by injecting electrical charge into its floating gate; the memory cell MC can be programmed to multiple levels, each one associated with a correspondingly range of its threshold voltage. Each level represents a different logical value; for example, the memory device 100 operates at 4 levels, so that each memory cell MC stores a logical value consisting of 2 bits (11, 10, 01 and 00 for increasing threshold voltages). Conversely, the memory cells MC are erased by removing the electrical charges trapped in their floating gates.

[0027] In the example at issue, the matrix 105m has a so-called NOR architecture. Particularly, the drain terminals of the memory cells MC of each column are connected to a corresponding bit line MBL, whereas the con-
trol gate terminals of the memory cells MC of each row are connected to a corresponding word line MWL; the source terminals of all the memory cells MC are connect-

e to a common terminal providing a reference voltage (or ground) GND. All the memory cells MC of the matrix 105m must be erased at the same time. On the other hand, the memory device 100 simultaneously reads/programs a word (for example, consisting of 16 bits). The bits of each word are stored in a page of memory cells MC (16/2=8 in the example at issue), which are associated with a single word line MWL; the bit lines MBL of the matrix 105m are grouped into 8 packets, each one associated with a different memory cell MC of every page.

[0028] The memory device 100 further includes a (single row) matrix 105r of reference cells RC, which are used for discriminating the different conditions of the memory cells MC. Each reference cell RC consists of a floating gate MOS transistor (substantially the same as the ones of the memory cells MC) with a predefined threshold voltage (typically set at the manufacturing of the memory device 100). The drain terminal of each reference cell RC is connected to a corresponding bit line RBL; the control gate terminals of all the reference cells RC are connected to a common word line RWL, whereas their source terminals are connected to the ground terminal GND.

[0029] The memory device 100 receives an address ADR for accessing a corresponding page. A portion of the address ADR is supplied to a row decoder 110, which selects the desired word line MWL. Another portion of the address ADR is supplied to a column decoder 115, which selects the desired bit line MBL in each packet. The memory device 100 further includes a ramp generator 120; the ramp generator 120 provides a biasing voltage Vb having a ramp-like pattern (i.e., with its value that increases linearly over time according to a constant slope). The ramp generator 120 may be of the analogue type (wherein the voltage of the biasing ramp Vb rises continuously) or of the digital type (wherein the voltage of the biasing ramp Vb rises at discrete intervals, with a step-like pattern approximating the desired ramp); in the latter case, the ramp generator 120 is generally implemented by means of a counter (incremented at a predefined frequency), which counter drives a Digital-to-Analog Converter (DAC).

[0030] The row decoder 110 couples the selected word line MWL with the ramp generator 120, so as to apply the biasing ramp Vb to the control gate terminals of the selected memory cells MC. The same biasing ramp Vb is also applied to the word line RWL through a driver 125 (which emulates the same loading conditions of the row decoder 110); in this way, the control gate terminals of all the reference cells RC receive the biasing ramp Vb in the same conditions as the selected memory cells MC.

[0031] The column decoder 115 connects the selected bit lines MBL to a control unit 130; the control unit 130 is also coupled with each bit line RBL through a correspond-

[0032] Considering now Figure 2, during a read operation the (unknown) threshold voltage of each selected memory cell MC (denoted with Vth) is compared with the ones of three (read) reference cells, denoted with RCri (i=0-2). The threshold voltage of each read reference cell RCri is equal to a read reference Vri comprised between the levels of the logical values i+1 and i; more specifically, the read reference Vr1 is comprised between the levels of the logical values 11 and 10, the read reference Vr2 is comprised between the levels of the logical values 10 and 01, and the read reference Vr0 is comprised between the levels of the logical values 01 and 00.

[0033] The biasing ramp Vb increases over time (t) from the ground voltage GND (0V) to a maximum value (denoted with Vb max) exceeding the highest read reference Vr2; for example, the biasing ramp Vb has a slope of 100.106 V/s and a length of 30-40ms, so as to reach the maximum value Vb max=3-4V. The biasing ramp Vb then remains at the maximum value Vb max for a further period required to complete the read operation (up to a time te, of the order of 50-60 ns).

[0034] When the biasing ramp Vb reaches the threshold voltage of the first read reference cell RCr1 (i.e., the read reference Vr1) at a time tr1, this read reference cell RCr1 turns on; the other read reference cells RCr2 and RCr0 turn on in succession when the biasing ramp Vb reaches the read reference Vr2 at a time tr2 and the read reference Vr0 at a time tr0, respectively.

[0035] Likewise, the (selected) memory cell MC turns on when the biasing ramp Vb reaches its threshold voltage Vth. Therefore, if this time (denoted with tm1) precedes the time tr2, the threshold voltage Vth is lower than the read reference Vr2; this means that the memory cell MC stores the logical value 1. Instead, if the turn on time of the memory cell MC (denoted with tm2) follows the time tr2, the threshold voltage Vth is higher than the read reference Vr2; this means that the memory cell MC stores the logical value 0. At the end, if the turn on time of the memory cell MC (denoted with tm1) follows the time tr2, the threshold voltage Vth is lower than the read reference Vr2 and the memory cell MC stores the logical value 0. At the end, if the turn on time of the memory cell MC (denoted with tm1) follows the time tr2, the threshold voltage Vth is lower than the read reference Vr2 and the memory cell MC stores the logical value 0.

[0036] The same procedure can also be used to verify the condition of the memory cell MC. Indeed, as shown in Figure 3a, whenever the memory cell MC is programmed to any logical value 10, 01 or 00, its threshold voltage Vth is brought over a program-verify reference Vv slightly higher than the corresponding read reference
Therefore, when the memory cell MC stores the i-th logical value its threshold voltage Vth is distributed over the verify reference VVr, below the read reference VR[i], of the next logical level (with the threshold voltage Vth that is distributed below the read reference VR2 for the logical value 11). In this way, the ranges of the threshold voltage Vth of the memory cell MC associated with the different logical values 11-00 are spaced apart; this facilitates the identification of the level of the threshold voltage Vth and then of the logical value stored in the memory cells MC, so as to increase the reliability of the read operation.

[0037] When the memory cell MC looses the electrical charge trapped in the floating gate, its threshold voltage Vth decreases accordingly. As a result, should the threshold voltage Vth fall below the read reference VR of the logical value originally stored in the memory cell MC, the threshold voltage Vth would enter the range of the lower level so as to cause the reading of a wrong logical value.

[0038] In order to prevent this situation, a guard reference Vg is associated with each read reference VR; the guard reference Vg is comprised between the read reference VR and the verify reference VVr, and defines a safe margin for the threshold voltage Vth. In this way, it is possible to detect a warning condition of the memory cell MC when its threshold voltage Vth reaches the guard reference Vg of the logical value stored therein (meaning that the memory cell MC is approaching the read reference VR because of the loss of electrical charge).

[0039] For this purpose, as illustrated in Figure 3b, during the above-described read operation the threshold voltage Vth of the memory cell MC is further compared with the ones of three (guard) reference cells, denoted with RCG; the threshold voltage of each guard reference cell RCG is equal to the guard reference Vg.

[0040] As above, when the biasing ramp Vb reaches the threshold voltage of the guard reference cell RCG (i.e., the guard reference Vg) at a time tg, this guard reference cell RCG turns on; the other guard reference cells RCG and RCG turn on in succession when the biasing ramp Vb reaches the guard reference Vg at a time t (following the time tr) and the guard reference Vg at a time t (following the time tr), respectively.

[0041] Therefore, if the time (denoted with tm) when the memory cell MC turns on follows a generic time tr but precedes the time t, its threshold voltage Vth is comprised between the read reference VR and the guard reference Vg; this means that the memory cell MC still stores the i-th logical value, but it has entered the corresponding warning condition.

[0042] In the solution according to an embodiment of the invention, the memory cell MC (in the warning condition) is then subjected to a refresh operation. For this purpose, a soft-program pulse Vp is applied to the memory cell MC (and more specifically to its drain terminal), so as to inject a small amount of electrical charge into the corresponding floating gate. The soft-program pulse Vp starts after the occurrence of the event that allows determining the condition of the memory cell MC (i.e., the entering the corresponding warning condition); this event consists of the turn on of the guard reference cell RCg (when the memory cell MC has turned on after the corresponding read reference cell RCg). The soft-program pulse Vp is applied (to the drain terminal of the memory cell MC) while it receives the biasing ramp Vb (at its control gate terminal). The soft-program pulse Vp has a value (for example, 4-5V) so that its combined action with the biasing ramp Vb slightly increases the threshold voltage Vth of the memory cell MC (for example, by 5-20mV). In this way, the threshold voltage Vth of the memory cell MC returns over the guard reference Vg associated with its logical value (thereby leaving the corresponding warning condition), or at least moves towards the guard reference Vg (away from the corresponding read reference VR).

[0043] The proposed solution avoids any waste of time for the refresh operation. Indeed, the time that is required in any case to complete each read operation is now exploited to refresh the same memory cells being read. In this way, the memory device implements a “shadow” refreshing, which is completely opaque to its operation.

[0044] As a result, it is possible to refresh the memory cells without having to stop operation of the memory device. At the same time, the refresh operation does not have any impact on the access time of the memory device.

[0045] All of the above fosters the widespread application of the refresh operation; this has a beneficial impact on the reliability of the memory device.

[0046] It should also be noted that the pattern of the biasing ramp Vb automatically ensures that the electrical field at the oxide layer of the memory cells MC is kept substantially constant while the electrical charge in their floating gates increases (during the refresh operation). Moreover, in this way the desired result is achieved with very low additional power consumption (being the biasing ramp Vb already available).

[0047] The soft-program pulse Vp may be generated with different techniques. With reference in particular Figure 4a, in an embodiment of the invention the soft-program pulse Vp is applied to the memory cell MC immediately (after the determination of its condition) up to the completion of the read operation. For example, when the memory cell MC stores the logical value 10 but has entered the corresponding warning condition (case A), the soft-program pulse Vp lasts from the time t (when the guard reference cell RCg2 turns on) to the time te (when the biasing ramp Vb ends). Likewise, when the memory cell MC is in the warning condition of the logical value 01 (case B), the soft-program pulse Vp lasts from the time t (when the guard reference cell RCg2 turns on) to the time te, whereas when the memory cell MC is in the warning condition of the logical value 00 (case C), the soft-program pulse Vp lasts from the time t to the time te. The figure also shows a voltage overdrive $\Delta V$ at
the control gate terminal of the memory cell MC (defined as the difference between the biasing ramp Vb applied thereto and its threshold voltage Vth).

[0048] The proposed implementation exploits the time available during the read operation to the utmost. However, the memory cells MC with the higher threshold voltages Vth receive the soft-program pulse Vp for a shorter period and with a lower voltage overdrive \( \Delta V \) (so that the effectiveness of the refresh operation is nonuniform).

[0049] Moving to Figure 4b, in a different embodiment of the invention the soft-program pulse Vp is still applied to the memory cell MC immediately (after the determination of its condition), but for a predetermined period Tp. More specifically, in all the cases described in the foregoing (A, B and C) the soft-program pulse Vp lasts from the time t0 to the time t0 +Tp. For this purpose, the period Tp must be set to allow the completion of the soft-program pulse Vp in every condition (i.e., even for the memory cells MC with the highest threshold voltages Vth); therefore, the period Tp is at most equal to Tp=te-tg0. As above, the figure also shows the corresponding voltage overdrive \( \Delta V \) at the control gate terminal of the memory cell MC.

[0050] In this implementation, the effectiveness of the refresh operation is substantially uniform. Indeed, both the length of the soft-program pulse Vp and the voltage overdrive \( \Delta V \) (disregarding the difference due to the final steady portion of the biasing ramp Vb) are independent of the threshold voltages Vth of the memory cells MC. However, the time available during the read operation is now exploited only partially.

[0051] In a further embodiment of the invention, as shown in Figure 4c, the soft-program pulse Vp is now applied at the same time to all the memory cells MC for the same period Tp. More specifically, in all the cases described in the foregoing (A, B and C) the soft-program pulse Vp lasts from the time t0 (when the last guard reference cell RCr0 turns on) to the time t0 +Tp, and preferably up to the completion of the read operation (i.e., to the time te). As above, the figure also shows the corresponding voltage overdrive \( \Delta V \) at the control gate terminal of the memory cell MC.

[0052] In the proposed implementation, the length of the soft-program pulse Vp is still independent of the threshold voltages Vth of the memory cells MC; the voltage overdrive \( \Delta V \) is now the highest possible in any condition (even if it is not uniform for the different threshold voltages Vth). As above, this result is achieved with a reduction of the time available during the read operation that is exploited.

[0053] Moving now to Figure 5, the control unit 130 is formed by the same structure repeated for each packet of bit lines MBL. Particularly, this structure includes a sense amplifier 505m for the (selected) memory cell MC, a sense amplifier 505r for each read reference cell RCr, and a sense amplifier 505g for each guard reference cell RCg, (which sense amplifiers 505m, 505r, and 505g bias the cells MC, RCr, and RCg to a common read voltage, such as 1V). The non-inverting input terminal of the sense amplifier 505m receives a current Im flowing through the selected bit line MBL (and then through the selected memory cell MC). The non-inverting input terminal of each sense amplifier 505r receives a current Ir flowing through the bit line (denoted with RBLr) associated with the read reference cell RCr; likewise, the non-inverting input terminal of each sense amplifier 505g receives a current Ig flowing through the bit line (denoted with RBLg) associated with the guard reference cell RCg. The inverting input terminals of all the sense amplifiers 505m, 505r, and 505g instead receive a comparison current Ic; the comparison current Ic is provided by a direct current generator 510 (referred to the ground terminal GND), and has a value lower than the current flowing through every memory/reference cell MC, RCr, and RCg, when conductive (for example, about 5-20 \( \mu A \)).

[0054] The sensing bits Sr (from the sense amplifiers 505r) are provided to an encoder 515; the encoder 515 converts the sensing bits Sr into a running value DTc, which represents the current logical value that would be stored in the memory cell MC at every instant during the read operation. Particularly (as shown in Figure 2), at the beginning of the read operation all the sensing bits Sr0-Sr2 are deasserted, and the running value DTc is equal to 11; the running value DTc then switches to 10, 01 and 00 at the assertion of the sensing bit Sr2 (time \( t_{r2} \)), of the sensing bit Sr1 (time \( t_{r1} \)) and of the sensing bit Sr0 (time \( t_{r0} \)), respectively. Preferably, the encoder 515 also latches each sensing bit Sr in this way, since the turn on of the corresponding read reference cell RCr has been detected (i.e., the sensing bit Sr has been asserted), this reference cell RCr can be turned off by means of the respective switch (so as to reduce the power consumption of the memory device and any noise caused by the current drained from the common ground terminal GND). The running value DTc is saved into a latch 520, in response to the assertion of the sensing bit Sm; in this way, the latch 520 provides the logical value actually stored in the memory cell MC (denoted with DT). As above, once the turn on of the memory cell MC has been detected (i.e., the sensing bit Sm has been asserted), the memory cell MC can be turned off by means of the column decoder; however, the biasing ramp Vb must be maintained up to the end of the read operation to detect the condition of the other memory cells CM in the selected page.

[0055] On the other hand, the sensing bits Sg (from the sense amplifiers 505g) are provided to a pulser 525, which also receives the Least Significant Bit (LSB) of the running value DTc (from the encoder 515). As described in detail in the following, the pulser 525 generates a guard pulse Pg that lasts from any toggle of the bit LSB (when
the running value DTc changes following the turn on of
a next reference cell RCr
i up to the assertion of the corresponding
sensing bit Sgi, (when the guard cell RCr
i turns on as well); in other words, the guard pulse Pg is
asserted when the memory cell MC would be in any warn-
ing condition at every instant during the read operation.
Preferably, the pulser 525 also latches each sensing bit
Sgi, so as to allow turning off the corresponding guard
reference cell RCr
i once its turn on has been detected.
The guard pulse Pg is saved into a latch 530, in response
to the assertion of the sensing bit Sm. The latch 530
outputs a guard strobe STRg, which is then asserted only
when the memory cell MC has turned on between the
turn on of a generic read reference cell RCr
i and of the
corresponding guard reference cell RCr
j (i.e., it is in the
respective warning condition); in this case, a refresh op-
eration must be performed on the memory cell MC.

For this purpose, the guard strobe STRg is input
to an AND gate 535; the AND gate 535 also receives a
signal ΔT, which is asserted for the desired length of the
soft-program pulse Vp to be applied to the memory cell
MC. The length signal ΔT is provided by a multiplexer
540, which receives three different signals (specifying
the length of the soft-program pulse Vp according to the
embodiments of the invention described in the foregoing
reference to Figures 4a-4c); the multiplexer 540
transmits one of them (so as to define the length signal
ΔT), in response to a selection signal SEL that is set to
the desired mode of operation of the memory device (for
example, stored in a configuration register).

Particularly, when the soft-program pulse Vp
must be applied immediately up to the completion of the
read operation (see Figure 4a), the multiplexer 540 trans-
mits a read signal RD that is asserted for the whole read
operation. On the other hand, when the soft-program
pulse Vp must be still applied immediately but for the
predetermined period Tp (see Figure 4b), the multiplexer
540 transmits a signal denoted with the same reference
(Tp), which is asserted for the length of that period; the
period signal Tp is generated by a pulser 545, which as-
serts the period signal Tp (for the desired time) in re-
sponse to the assertion of the guard strobe STRg. At the
end, when the soft-program pulse Vp must be applied at
the end of the read operation (see Figure 4c), the multi-
plexer 540 transmits a top signal HI that is output by a
NOR gate 550; the NOR gate 550 receives the two bits
of the running value DTc (denoted with DTc0 and DTc1),
so that the top signal HI is asserted when the last read
reference cell RCr
i turns on (running value DTc=00) until
the completion of the read operation.

In any case, the AND gate 535 outputs an en-
abling signal EN, which is asserted for the length defined
by the signal ΔT whenever the threshold voltage of the
memory cell MC is in any warning condition (guard strobe
STRg asserted). The signal EN enables a soft-program
load 555. While the enabling signal EN is asserted, the
soft-program load 555 applies the soft-program pulse Vp
to the bit line MBL (and then to the drain terminal of the
memory cell MC). For example, the soft-program load
555 connects the bit line MBL to a charge pump (not
shown in the figure), which generates the high voltage
of the soft-program pulse Vp (4-6V) from a lower power
supply voltage +Vdd of the memory device (for example,
2-3V); it should be noted that the charge pump typically
has a driving capability far higher than the one of the
sense amplifier 505m, so as to force the desired voltage
at the bit line MBL.

With reference now to Figure 6, the pulser 525
of the above-described control unit includes a stage for
each sensing bit Sgi. The generic stage has a right branch
formed by two NMOS transistors 605R
i and 610R
i and
a left branch formed by two NMOS transistors 605L
i and
610L
i. Particularly, the gate terminals of the transis-
tors 605R
i and 605L
i receive the (latched) sensing bit Sgi,
whereas the gate terminals of the transistors 610R
i and
610L
i receive its inverted value (denoted with RSgi)

through an inverter 615
i. The drain terminals of the trans-
stors 605R
i and 610R
i, 605L
i and 610L
i, are connected together to define
a node NR
i, whereas the drain terminals of the transis-
tors 605L
i and 610L
i are connected together to define a node
NL
i. In each stage different from the first one (i>0), the
source terminal of the transistor 605R
i is connected to
the node NR
i-1 whereas the source terminal of the transis-
tor 605L
i is connected to the ground terminal GND and
the source terminal of the transistor 605L
i is connected
to a terminal providing the power supply voltage+
Vdd. In each even stage (i.e., i=0,2 in the example at
issue), the source terminal of the transistor 610R
i receives
the power supply voltage +Vdd and the source termi-
nal of the transistor 610R
i receives the ground volt-
age GND; conversely, in each odd stage (i.e., i=1 in the
example at issue), the source terminal of the transistor
610R
i receives the ground voltage GND and the source
terminal of the transistor 610R
i receives the power supply voltage +Vdd.

The pulser 525 further includes a latch formed
by two PMOS transistors 620R and 620L. Particularly,
the transistor 620R has the drain terminal connected to
the node NR
i and the transistor 620L has the drain ter-
minal connected to the node NL
i of the last stage. The
gate terminal of the transistor 620R is connected to the
drain terminal of the transistor 620L, and the gate termi-
nal of the transistor 620L is connected to the drain termi-
nal of the transistor 620R. The source terminals of both
the transistors 620R and 620L receive the power supply volt-
age +Vdd.

The drain terminal of the transistor 620R provides
a selection signal SEL, and the drain terminal of the
transistor 620L provides its inverted value (denoted with
SEL). The selection signal SEL controls a pass gate
625R that receives the bit LSB, whereas the inverted
selection signal SEL controls a pass gate 625L that re-
ceives its inverted value (denoted with LSB) through an
inverter 630. The output terminals of the two pass gates
625R and 625L are connected together to provide the guard pulse Pg.

[0062] Considering now Figure 6 together with Figure 3b, at the beginning of the read operation the bit LSB is deasserted (since the running value DTC is 00), and then the inverted bit LSB is asserted; at the same time, all the sensing bits Sg are deasserted. In this condition, all the transistors 605R, 605L are off, while all the transistors 610R, 610L are on (disregarding their threshold voltages). Therefore, in each even stage (i.e., i=0,2) the node NL is at the power supply voltage +Vdd and the node NL is at the ground voltage GND, whereas in each odd stage (i.e., i=1), the node NR is at the ground voltage GND and the node NL is at the power supply voltage +Vdd; as a result, the selection signal SEL is asserted (node NR at the power supply voltage +Vdd) and the inverted selection signal SEL is deasserted (node NL at the ground voltage GND). The pass gate 625R is then enabled to transfer the bit LSB (while the pass gate 625L is disabled), so that the guard pulse Pg is deasserted.

[0063] As soon as the running value DTC changes (time t2), the bit LSB toggles so that the bit LSB becomes asserted and the inverted bit LSB become deasserted; as a result, the guard pulse Pg is automatically asserted. Later on, the sensing bit Sg is asserted as well (time t2); therefore, the transistors 605R, 605L turn on and the transistors 605L, 605R turn off. The node NR is then connected to the node NR, and the node NL is connected to the node NL; therefore, from now on the stage associated with the sensing bit Sg becomes completely opaque to the operation of the pulsers 525. In this condition, the selection signal SEL is deasserted (node NR at the ground voltage GND) and the inverted selection signal SEL is asserted (node NL at the power supply voltage +Vdd). The pass gate 625L is then enabled (instead of the pass gate 625R), now transferring the inverted bit LSB that deasserts the guard pulse Pg. The same operations described above are repeated at any next changes of the running value DTC and at the assertions of the corresponding sensing bits Sgi.

[0064] An alternative implementation of the control unit of the memory device is illustrated in Figure 7 (this implementation is differentiated by adding an apex to its reference, i.e., 130'); for the sake of simplicity, the elements corresponding to the ones shown in the Figure 5 are denoted with the same references and their explanation will be omitted.

[0065] In this case, the memory device does not include any guard reference cell; on the contrary, the same function is emulated by exploiting the read reference cells RCr themselves. For this purpose, the guard pulse Pg (from a different pulsers denoted with 525') is also applied to a delay block 705. The delay block 705 returns the same signal (denoted with DPg) to the pulsers 525 with a predetermined delay. The delay is set to a value (for example, about 0.2-2 ns) corresponding to a voltage difference along the biasing ramp Vb of 20-200 mV (assuming a slope of 100·10⁸ V/s); preferably, trimming elements (not shown in the figure) are used to regulate this delay to the desired value.

[0066] As a result, the delayed guard pulse DPg is asserted (with its delay) after each assertion of the guard pulse Pg, i.e., after the assertion of each sensing bit Sr indicating the turn on of the corresponding read reference cell RCr. In other words, the assertion of the delayed guard pulse DPg occurs when the biasing ramp Vb has reached a value slightly higher than the threshold voltage of the read reference cell RCr (i.e., the corresponding read reference VR); therefore, the assertion of the delayed guard pulse DPg emulates the turn on of another reference cell with a slightly higher threshold voltage, which can be set to the desired guard reference VG (trimming the delay accordingly). This technique is described in greater detail in the co-pending European patent application No.EP05101680.8 filed on 3 March 2005 (the entire disclosure of which is herein incorporated by reference).

[0067] The pulsers 525' asserts the guard pulse Pg whenever the bit LSB toggles (when the running value DTC changes following the turn on of a next read reference cell RCr) up to the assertion of the delayed guard pulse DPg (emulating the turn on of the corresponding guard reference cell). Therefore, the operation of the control unit 130' is exactly the same as the one described above with reference to Figure 5.

[0068] The proposed implementation simplifies the structure of the memory device (since the number of reference cells is now reduced). This has a beneficial impact on the testing time required to trim the reference cells. Moreover, the difference between the read references and the corresponding guard references can be defined with a very high accuracy (since any change of the threshold voltages of the read reference cells only shifts the (read/guard) references, but it is completely immaterial to their differences.

[0069] A possible structure of the pulsers 525' suitable to be used in the above-described control unit is shown in Figure 8; as above, the elements corresponding to the ones shown in Figure 6 are denoted with the same references and their explanation will be omitted.

[0070] Particularly, the pulsers 525' now includes a (single-stage) right branch formed by two NMOS transistors 705R and 710R, and a left branch formed by two NMOS transistors 705L and 710L. Particularly, the gate terminal of the transistor 705R receives the bit LSB, whereas the gate terminal of the transistor 705L receives the inverted bit LSB; the source terminals of both the transistors 705R and 705L receives the ground voltage GND. The drain terminal of the transistor 705R is connected to the source terminal of the transistor 710R to define a node NNa, and the drain terminal of the transistor 705L is connected to the source terminal of the transistor 710L to define a node NLa. The gate terminals of both the transistors 710R and 710L receive the delayed guard pulse DPg. The drain terminal of the transistor 710R defines a node NRb that is connected to the drain terminal of the transistor 620R,
and the drain terminal of the transistor 710L defines a node NLb that is connected to the drain terminal of the transistor 620L (with the remaining part of the pulser 525' that has the same structure described above). The pulser 525' further includes another NMOS transistor 715. The transistor 715 has the source terminal that receives the ground voltage GND, and the drain terminal that is connected to the node NLb; the gate terminal of the transistor 715 receives a reset signal RST.

[0071] At the beginning of the read operation, the reset signal RST is asserted for a short period; in this way, the transistor 715 turns off and brought the node NLb to the ground voltage GND. As a result, the transistor 620R turns on (selection signal SEL' asserted), while the transistor 620L remains off (inverted selection signal SEL' deasserted). At the same time, the bit LSB is deasserted and the inverted bit LSB is asserted; therefore, the pass gate 625R is enabled to transfer the bit LSB to the transistor 705L (while the pass gate 625L is disabled), so that the guard pulse Pg is deasserted. Moreover, the transistor 705R is off and the transistor 705L is on (bringing the node NLa to the ground voltage GND). The delayed guard pulse DPg it deasserted as well; therefore, the transistors 710R and 710L are off so as not to affect the condition of the latch 620R, 620L.

[0072] Whenever the running value DTc changes (because a generic read reference cell RCr turns on) the bit LSB toggles, so that the bit LSB becomes asserted and the inverted bit LSB become deasserted; as a result, the guard pulse Pg is automatically asserted. At the same time, the transistor 705R turns on (bringing the node NRa to the ground voltage GND) and the transistor 705L turns off. The delayed guard pulse DPg is asserted as well after the desired delay, so as to turn on the transistors 710R and 710L. In this condition, the node NRb is connected to the node NRa, and it is brought to the ground voltage GND. As a result, the transistor 620L turns on and the inverted selection signal SEL' is asserted, while the transistor 620R turns off and the selection signal SEL' is deasserted. The pass gate 625L is then enabled (instead of the pass gate 625R), now transferring the inverted bit LSB that deasserts the guard pulse Pg. The same operations described above are repeated at any next changes of the running value DTc.

[0073] Naturally, in order to satisfy local and specific requirements, a person skilled in the art may apply to the solution described above many modifications and adaptations. Particularly, although the present invention has been described with a certain degree of particularity with reference to preferred embodiment(s) thereof, it should be understood that various omissions, substitutions and changes in the form and details as well as other embodiments are possible; moreover, it is expressly intended that specific elements and/or method steps described in connection with any disclosed embodiment of the invention may be incorporated in any other embodiment as a general matter of design choice.

[0074] For example, the numerical examples described above are merely illustrative and must not be interpreted in a limitative manner.

[0075] Similar considerations apply if the memory device has another structure or includes equivalent units. Particularly, the application of the proposed solution to different non-volatile memory devices (such as EPROMs, flash memories with NAND architecture, phase-change memories), or even to volatile memory devices (such as DRAMs) is contemplated. Moreover, the modification of the above-described circuits for a memory device operating with a different number of levels (even not a power of 2) is straightforward; in any case, the technical idea of the invention may also be applied to a standard memory device, wherein each memory cell stores a single bit.

[0076] Alternatively, the biasing ramp may be generated in another way (for example, by applying a direct current to the selected memory cells and to the reference cells); similar considerations apply if the biasing voltage has a different monotone time pattern (either increasing or decreasing over time). Moreover, the values stored in the memory cells can be determined with equivalent techniques based on the temporal order of their turning on with respect to the ones of the reference cells (for example, by latching an indication of the time at which each memory/reference cell turns on).

[0077] It is emphasized that the proposed solution is suitable to be applied to any reading of the memory cells. For example, the same technique may be implemented during a program-verify operation (when the memory cells are read to determine whether the desired logical values have been correctly written). Alternatively, it is also contemplated the possibility of reading all the memory cells (with their refreshing) in response to a dedicated command; for example, this command may be invoked at the power-on of the memory device (preferably disabling the output of the logical values read from the memory cells).

[0078] Similar considerations apply if the soft-program pulse has a different value and/or shape. In any case, although in the preceding description great emphasis has been given to a refresh operation based on a soft-program pulse, it should be noted that the proposed solution has a more general applicability to any writing of the memory cells; particularly, if the memory device supports the erasure at the level of single memory cells, it is also possible to verify whether electrical charge has been injected into their floating gates (with an undesired programming thereof), and then apply a soft-erase pulse when it is necessary.

[0079] In a different embodiment of the invention, the refresh operation is prevented (for a predetermined period) at the power-on of the memory device (when the required high voltage provided by the charge pump may be not available yet); however, the generation of the soft-program pulse from the power supply voltage directly is not excluded (taking into account that the soft-program pulse generally has a lower value than the very high one...
required during an actual programming of the memory cells). Moreover, nothing prevents making the guard strobe available outside the memory device to signal the occurrence of any warning condition.  

[0080] Without departing from the principles of the invention, it is also possible to implement a blind refresh operation (without any verification of the warning condition of the memory cells), for example, by exploiting a soft-program pulse that is shorter and/or lower (especially in standard memory devices wherein no risk exists of writing a wrong logical value by reaching the next level of the threshold voltage).

[0081] Similar considerations apply if the entering of the warning conditions (in the implementation based on the guard reference cells) is detected in a similar way or by means of equivalent circuits.  

[0082] Likewise, the emulation of the guard reference cells may be obtained with alternative techniques (for example, by increasing/decreasing the latched time at which each memory and/or reference cell turns on); in this case as well, equivalent circuits may be exploited.  

[0083] Even though in the preceding description reference has been made to specific techniques for generating the soft-program pulse, they are not to be interpreted in a limiting manner. For example, in a different embodiment of the invention it is possible to adapt the value of the soft-program pulse according to the actual condition of the memory cell (such as increasing it for the higher threshold voltages to compensate the reduced effectiveness of the refresh operation). Alternatively, the soft-program pulse may be maintained after the completion of the biasing ramp (such as up to the submission of a next command for accessing the memory device); in this case, it would also be possible to verify the condition of the memory cell after the refresh operation. Of course, it not strictly necessary for the memory device to support all the proposed techniques at the same time (with the structure of the control unit that is simplified accordingly when a single one of them is implemented).

[0084] It should be readily apparent that the proposed solution may be part of the design of an integrated circuit. The design may also be created in a programming language; moreover, if the designer does not fabricate chips or masks, the design may be transmitted by physical means to others. In any case, the resulting integrated circuit may be distributed by its manufacturer in raw wafer form, as a bare die, or in packages form. Moreover, the circuit may be integrated with other chips in intermediate products (such as mother boards). In any case, the integrated circuit is suitable to be used in complex systems (such as mobile phones).

Claims

1. A refresh circuit (130) for refreshing a memory device (100), the refresh circuit including:

   a. reading means (110-125,505m,505r,515-520) for reading a set of memory cells (MC), the reading means including means (110-125) for applying a biasing voltage having a substantially monotone time pattern to the memory cells and to a set of reference cells (RC) each one having a predefined reference threshold voltage, means (505m,505r) for detecting the reaching of a comparison current (Ic) by a cell current (Im) of each memory cell and by a reference current (Ir) of each reference cell, and means (515-520) for determining a condition of each memory cell according to a temporal relation of the reaching of the comparison current by the corresponding cell current and the reference currents, and

   b. writing means (555) for applying a writing voltage to at least one selected of the memory cells, characterized in that

   c. the refresh circuit further includes control means (505g,525-550;525',705,530-550) for enabling the writing means during at least part of the application of the biasing voltage after the determination of the condition of each selected memory cell.

2. The refresh circuit (130) according to claim 1, wherein in the control means (505g,525-550;525',705,530-550) includes verifying means (505g,525-530;525'-530,705) for detecting a warning condition of each memory cell when the difference between the corresponding threshold voltage and any reference threshold voltage is lower than a predefined safe margin, the memory cells (MC) being selected for the application of the writing voltage in response to the detection of the warning condition.

3. The refresh circuit (130) according to claim 2, wherein in the means for reading further includes means (110-125) for applying the biasing voltage to a set of further reference cells (RC) each one having a predefined further reference threshold voltage, the difference between each further reference threshold voltage and a corresponding reference threshold voltage being equal to the safe margin, and wherein the verifying means (505g,525-530) includes means (505g) for detecting the reaching of the comparison current (Ic) by a further reference current (Ir) of each further reference cell, and means (525-530) for determining the warning condition of each memory cell (MC) according to a temporal relation of the reaching of the comparison current by the corresponding cell current (Ic), the reference currents (Ir) and the further reference currents.

4. The refresh circuit (130) according to claim 3, wherein in the means (525-530) for determining the warning condition includes means (525) for providing a guard signal (Pg) being asserted when the biasing voltage
The refresh circuit (130) according to any claim from 6.

The refresh circuit (130) according to claim 5, the means (525) for providing the guard signal including means (625R,625L) for setting the guard signal alternatively to a bit changing between each pair of consecutive logical values stored in the memory cell or to said bit being inverted, and means (605R, 605L, 610R, 610L, 615R, 615L, 620R, 620L) for exchanging the bit and the inverted bit in response to the detection of the reaching of the comparison current (Ic) by each further reference current (Ir).

5. The refresh circuit (130) according to claim 2, wherein in the verifying means (525',705,530-550) includes means (705) for delaying the detection of the reaching of the comparison current (Ic) by each reference current (Ir) according to a predefined delay, the change of the biasing voltage (Vb) corresponding to the delay being equal to the safe margin, and means (525',530) for determining the warning condition of each memory cell (MC) according to a temporal relation of the reaching of the comparison current (Ic) by the corresponding cell current (Ic), the reaching of the comparison current by the reference currents (Ir), and the delayed detections.

6. The refresh circuit (130) according to claim 5, the means (525',530) for determining the warning condition includes means (525') for providing a guard signal (Pg) being asserted when the biasing voltage (Vb) is comprised between any reference threshold voltage and the corresponding further reference threshold voltage, the means (705) for delaying the detection including means for delaying the guard signal, and means (530) for latching the guard signal in response to the reaching of the comparison current (Ic) by the cell current (Im), the means (525') for providing the guard signal including means (625R,625L) for setting the guard signal alternatively to a bit changing between each pair of consecutive logical values stored in the memory cell or to said bit being inverted, and means (705R,705L,710R,710L,620R,620L) for exchanging the bit and the inverted bit in response to each assertion of the delayed guard signal (DPg).

7. The refresh circuit (130) according to any claim from 2 to 6, wherein the control means (505g,525-550; 525',705,530-550) includes means (540,545) for enabling the writing means (555) for a predefined time interval from the detection of the warning condition.

8. The refresh circuit (130) according to any claim from 2 to 7, wherein the control means (505g,525-550; 525',705,530-550) includes means (540,545) for enabling the writing means (555) for a predefined time interval from the detection of the warning condition.

9. The refresh circuit (130) according to any claim from 2 to 8, wherein the control means (505g,525-550; 525',705,530-550) includes means (540,550) for enabling the writing means (555) from the detection of the reaching of the comparison current by the reference current of a last reference cell having the highest threshold voltage in absolute value to a completion of the reading.

10. A memory device (100) including a plurality of memory cells (MC), means (110,115) for addressing a set of memory cells, a set of reference cells (RC) each one having a predefined reference threshold voltage, and the refresh circuit (130) according to any claim from 1 to 9.

11. A method of refreshing a memory device (100), the method including the steps of: reading (110-125,505m,505r,515-520) a set of memory cells (MC), the step of reading including applying (110-125) a biasing voltage having a substantially monotone time pattern to the memory cells and to a set of reference cells (RC) each one having a predefined reference threshold voltage, detecting (505m,505r) the reaching of a comparison current (Ic) by a cell current (Im) of each memory cell and by a reference current (Ir) of each reference cell, and determining (515-520) a condition of each memory cell according to a temporal relation of the reaching of the comparison current by the corresponding cell current and the reference currents, and applying (555) a writing voltage to at least one selected of the memory cells, characterized in that

the method further includes the step of:

enabling (505g,525-550;525',705,530-550) the application of the writing voltage during at least part of the application of the biasing voltage after the determination of the condition of each selected memory cell.
**DOCUMENTS CONSIDERED TO BE RELEVANT**

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document with indication, where appropriate, of relevant passages</th>
<th>Relevant to claim</th>
<th>CLASSIFICATION OF THE APPLICATION (IPC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>US 6 188 610 B1 (KAKIZOE KAZUHIKO ET AL) 13 February 2001 (2001-02-13) * column 6, line 44 - column 7, line 25; figure 5 *</td>
<td>1-11</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>EP 1 473 739 A (DIAGRAM SEMICONDUCTOR GMBH) 3 November 2004 (2004-11-03) * paragraphs [0022] - [0025]; figures 3,4 *</td>
<td>1-11</td>
<td></td>
</tr>
</tbody>
</table>

**TECHNICAL FIELDS SEARCHED (IPC)**

| G11C |

---

The present search report has been drawn up for all claims

**Place of search**

Munich

**Date of completion of the search**

13 September 2006

**Examiner**

Gaertner, Wolfgang
This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

13-09-2006

<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family member(s)</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td>US 2003048665 A1</td>
<td>13-03-2003</td>
<td>NONE</td>
<td></td>
</tr>
<tr>
<td>EP 1271552 A</td>
<td>02-01-2003</td>
<td>NONE</td>
<td></td>
</tr>
</tbody>
</table>

For more details about this annex: see Official Journal of the European Patent Office, No. 12/82
REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader’s convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- EP 1271552 A [0009]
- EP 05101660 A [0066]